OM nucleic - nucleic search, using sw model

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: May 3, 2003, 23:25:47 ; Search time 2710.34 Seconds
                                        (without alignments)
                                        17782.898 Million cell updates/sec

Title: US-10-027-000-1
Perfect score: 2976
Sequence: 1 ttatagtcgctttgttaaat........aaaaaaaaaaaaaaaaa 2976
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32208132

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estln:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_estom:*
                 17: gb_gss:*
                 18: em_gss_hum:*
                 19: em_gss_inv:*
                 20: em_gss_pln:*
                 21: em_gss_vrt:*
                 22: em_gss_fun:*
                 23: em_gss_mam:*
                 24: em_gss_mus:*
                 25: em_gss_other:*
                 26: em_gss_pro:*
                 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	290.8	9.8	1056	17	CNS078CQ	AL433824 T3 end of
c 2	254.6	8.6	503	17	AQ399568	AQ399568 mgxb0010D
c 3	227	7.6	1021	17	CNS0766A	AL431000 T7 end of
c 4	187.4	6.3	994	17	CNS0772N	AL433353 T3 end of
c 5	184.6	6.2	970	17	CNS075SU	AL430516 T3 end of
c 6	176.8	5.9	1045	17	CNS07387	AL427181 clone BA0
c 7	175.4	5.9	952	17	CNS073UC	AL427978 clone BA0
c 8	171.8	5.8	812	17	CNS072ML	AL426403 clone BA0
c 9	167.2	5.6	917	17	CNS070RO	AL423994 T3 end of
c 10	165.4	5.6	745	17	CNS072H1	AL426203 clone BA0
c 11	158.6	5.3	993	17	CNS078JO	AL434074 T3 end of
c 12	158.4	5.3	986	17	CNS076GU	AL431380 T3 end of
c 13	153.4	5.2	836	17	CNS074WR	AL429361 clone BA0
c 14	126.2	4.2	860	17	CNS0770T	AL432099 T3 end of
c 15	120.6	4.1	869	17	CNS075M9	AL431963 T7 end of
c 16	115.6	3.9	1010	17	CNS076X1	AL431457 T3 end of
c 17	115.4	3.9	633	9	CNS074X5	AL714457 AL714436
c 18	115	3.9	684	9		AL714436
c 19	109.2	3.7	521	17		AZ928908
c 20	105.2	3.5	984	12	CNS0700H	AZ928908 479 difi9
c 21	104.6	3.5	573	12	CNS07ZXT	BE805639 ss43f04.y
c 22	83.4	2.8	1000	17	CNS07ZXT	AL426807 clone BA0
c 23	82.2	2.8	1132	13		BM320864 rockefell
c 24	79.8	2.7	865	17	CNS07709	AL433015 T3 end of
c 25	79.2	2.7	376	9	AL513757	AL513757
c 26	76.8	2.6	610	9	AL827475	AL827475 AL827475
c 27	75.4	2.5	781	17	CNS072EL	AL425827 clone BA0
c 28	75.4	2.5	1012	17	CNS072L	AL433351 T3 end of
c 29	72.6	2.4	861	10	BE636696	BE636696 rockefel
c 30	72.2	2.4	331	10	BE059900	BE059900 sn38f02.y
c 31	71.8	2.4	194	10	AW184930	AW184930 se83g11.y
c 32	71.6	2.4	997	14	BQ276691	BQ276691 AGENCOURT
c 33	71.6	2.4	1165	13	BM320900	BM320900 rockefell
c 34	71	2.4	1057	13	BM321375	BM321375 rockefell
c 35	70.8	2.4	226	12	BG792492	BG792492 UTSW_H38A
c 36	70.8	2.4	226	12	BG817252	BG817252 UTSW_H20A
c 37	70.8	2.4	486	9	AI297124	AI297124 LPI1346.5
c 38	70.4	2.4	231	10	AW346777	AW346777 29458 MAR
c 39	70.4	2.4	266	13	BI782468	BI782468 kh26b06.y
c 40	70.4	2.4	641	12	BG809574	BG809574 mgct001xa
c 41	70.4	2.4	824	10	BE040678	BE040678 OF04H06 O
c 42	70.2	2.4	705	17	BH019836	BH019836 L5336c.d_
c 43	70.2	2.4	732	10	BH018404	BH018404 L14841.d_
c 44	70.2	2.4	853	13	BM321393	BM321393 rockefell
c 45	70	2.4	358	9	AL515173	AL515173 AL515173

ALIGNMENTS

RESULT 1
CNS078CQ/c CNS078CQ   1056 bp   DNA   linear   GSS 08-JUL-2001
LOCUS   CNS078CQ
DEFINITION T3 end of clone BB0AA019F09 of library BB0AA from strain CBS 4732
            of Pichia angusta, genomic survey sequence.
ACCESSION   AL433824
VERSION   AL433824.1   GI:12217238
KEYWORDS   GSS.
SOURCE   Pichia angusta.
ORGANISM   Pichia angusta.
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1 (bases 1 to 1056)
AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvelise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE   20584711
PUBMED   11152876
REFERENCE   2 (bases 1 to 1056)
AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
TITLE   Genomic exploration of the hemiascomycetous yeasts: 13. Pichia
            angusta

```
JOURNAL     FEBS Lett. 487 (1), 76-81 (2000)
MEDLINE     20584723
PUBMED      11152888
REFERENCE   3 (bases 1 to 1056)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage;
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..1056
                     /organism="Pichia angusta"
                     /strain="CBS 4732"
                     /db_xref="taxon:4905"
                     /clone="BB0A019F09"
                     /clone_lib="BB0AA"
     misc_feature    /note="end : T3"
                     complement(<105..>1019)
                     /note="similar to P07337 ( Beta-glucosidase precursor BGLS
                     ] Kluyveromyces marxianus ]"
                     /evidence=not_experimental

BASE COUNT     245 a    290 c    258 g    259 t    4 others
ORIGIN

Query Match        9.8%;
Best Local Similarity  56.7%;  Score 290.8;  DB 17;  Length 1056;  Pred. No. 7.5e-30;
Matches 589;  Conservative 3;  Mismatches 441;  Indels 6;  Gaps 3;

Qy    414 GACTTCTTGGCACACAAAAGGCTCTCCCCAAGCAATGGAGTCCCCTCGTTTACAGAT 473
          |||| ||||| || || || ||||||||  |||  ||||| | || || |  |||||
Db   1044 GACTTACTGGCAATACGACGCGGTCTGCCACAGTTTGRAATTCCCCAAAAATTCGTACTTCGGAC 985

Qy    474 GGGCCCC-AACGGCGTAAGAGGGACCAAGTTCTTCAATGGCCGTCTCCCTGCGGCCTGCTTCCC 532
          ||||:   |||| | || ||   || ||  || || ||   |||| || ||||||  |||||||
Db    984 GGCCCMAAATGGTATTCGGTGGCCAGAACCACTTCAACAGTACTCCGACCCCATGTATCCC 925

Qy    533 TTGCGGCACGTCGCTCGGTTCCACATTCAACCAAACTCTCCTCGAAGAGCAGTAAGAT 592
          |||||  || |  || | ||  |||| || ||  ||| | |||||||| || || |||
Db    924 ATCAGGAACAGCCCTTCGCCTCCACTTGGGATGRATGACGANCTTTTGTTCCACTGGGCGTATT 865

Qy    593 GATGGGCAAAGAGGCCATCGGTCTAAGAGTGCCCATGTGATCCCTCCGGCCCCGACTATCAACAT 652
          ||||| ||||| |||||||||   || ||  |||||  ||||||  |||||| || |||   |||
Db    864 GTTGGACAAGGAGCCCGAGCCAAAGGCCGTRGCAGGTRGAGTCTCGGGCCCCAGCCAACAT 805

Qy    653 GCAACGCTCCCCTCCTCCGGTGGACGTGGCCTTCGACTGCGAGTTGGTAGGAGATCCGTTCCTGGCC 712
          ||||| |  || |||||  |||| |||   |||  || ||||||  |||||||  ||| ||||
Db    804 CCAACGTGTTCCTCCTTCGGCCGGACGTGGGTTTGAGTGCATTTTTCGGAAGATCCGTRTCTCTC 745

Qy    713 GGGCTTGGGACTCGCGGCTCTCATCCGCGGCCATTCAGACGCCGGAGTGCAGGCTACGAT 772
          |||||||  |      |    :||| |||||| ||| | ||    |||  ||| |||||
Db    744 GGGAATCCTGTCTGCGGCGTTCAAAAGTCAGGTGYTCAAAAGTCAGGTCTTGCCTGCCACCAT 685

Qy    773 CAAGCACTTTTGTGTGCAATGATCAGGAGGACAGCAGGCGCATGATGGGTGCAGAGCAGTCGTCAC 832
          ||||||||| |||   |||| ||   ||||| | || ||  || ||| ||| |||||||   ||
Db    684 GAAGCACTTTGTCGGCACAACGACCAGGAGCACGCGTTTCACGTTGCGAATTCAGTGAGTTC 625

Qy    833 GGAGCGGGCGTCTCCGTGAAATCTTACGCACTCCCGTTCCAGATTTGCTGTGCGAGACTCCCA 892
          | ||||||  |  ||  |||| || || || ||| ||||  ||||| | || ||||| ||
Db    624 GCAGCGGTGCCTTGCCAGAGATATACCTCCGTCCGTTTGAGATCGGCGATCAAGGAATCCA 565

Qy    893 GCCCGGTGCCTTCATGACGGCGTTACAATGGCATCAATGGCGTTGCTCGTGCACGGAGAACCC 952
          ||| |||  || |||  ||||||||| || || ||| | ||| || | || |  | ||||
Db    564 GCCCAAAGCTATTATGACAAGTCACGGAGTCACGTGACGCCATGTGACGCTAT 505
```

```
RESULT 2
AQ399568

LOCUS       AQ399568     503 bp     DNA     linear     GSS     06-MAR-1999
DEFINITION  mgxb00100D05f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
            clone mgxb00100D05f, DNA sequence.
ACCESSION   AQ399568
VERSION     AQ399568.1  GI:4370595
KEYWORDS    GSS.
SOURCE      Magnaporthe grisea.
  ORGANISM  Magnaporthe grisea
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE   1 (bases 1 to 503)
  AUTHORS   Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.;
            Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
            Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Dean RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson Universiy, Clemson, SC 29634
            Tel: 864 656 5737
            Fax: 864 656 4293
            Email: rdean@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence stop: 280.
FEATURES             Location/Qualifiers
     source          1..503
                     /organism="Magnaporthe grisea"
                     /strain="70-15"
                     /db_xref="taxon:148305"
                     /clone="mgxb00100D05f"
                     /clone_lib="CUGI Rice Blast BAC Library"

Qy    953 TAAATATCTTGATGGGAGTGCTTCGAAAGGAATGGGGGTTGGGAGTGGGGCTAATCATCAGAGCGA 1012
          || |||| ||| ||| || ||  ||  |||||  ||  || |||||  ||||||| || |||||
Db    504 ATACCTGCTCTCTGAGGTTCTCCGTCACGAGTGGGGCTATCAAGGCCTCATAAATGTCAGA 445

Qy   1013 CTGGTACGGCACATACAGTACCACACAGAAGCCGGTTGTTGGCAGGCGCCTCGACCTCAGATGCC 1072
          |||||  ||| || ||| || |  |||||  |||| |||| ||||  || ||||  |||||||
Db    444 CTGGCACAGCCACTTTCCACACAGAAGCCGGTTGCTGGCCTGGGACTTGAGTTGGCAGAATGCC 385

Qy   1073 CGGACCTCCACCGCTTCCGAGGAGAAACACTCAAGGTCAGATGCCTCCAACGGAAAGCCCTT 1132
          |||| |  ||  || ||||| || || |  || ||| |  | || || |||||||  ||||
Db    384 CGGCCCCGGAATATGGAGAGGAGGCTRAACCTCAGTAGGGCGGTTGCTGGCTGGGCAAAATTTC 325

Qy   1133 TATCCACGTCATTGACCAGACGGGGCTAGGGAAGTTCTTCAGTTCGTCAAGAAGTGTGCTGC 1192
          || ||| || || |||| || |  ||  || || || || || ||  |||| || || ||
Db    324 GGATGACGTATTGGACCAGCCGTCGTAAGGTGCTCCAGTTGCAATTGCGGAAACAGCGTCATAGA 265

Qy   1193 CTCCGGAGTGACGGACGAAACGGCCCCGAGACGACGACTGTCAACAACACCCCGAAAACGGAGC 1252
          | |||  || | ||  |||  ||| | |||||  |  |||||||||||||| |||||| ||
Db    264 TGCGGGAATTGAGGACACAATAAGGCAGAGGAG--CAGCCTGATACCCCTGAATACCCTGAAACAAGAGC 208

Qy   1253 TCTCCTCCGGAAGGTTGGCAACGAGGGCCATCGTGCTGCTGAAGAAACGACGAGACAACGGTTCT 1312
          || |||  ||||| ||| ||||| ||  |||||||||  || ||| | |  |  ||||||| |||
Db    207 GTTTTTGAGGAAAAATAGCTGGTGAGGCTGCGTGCGTCGTGCCTTCTCAAAAACAAGGAATGGGTGTGCT 148

Qy   1313 GCCCTTGAGCAAGAAGAAGAGAAGAGCGCTGATTGTCGGGCCCCAACGGCCAAGCAGCCCACATA 1372
          ||  || ||||| || |||   || || | ||| || |||  ||||||  ||||| ||  ||||
Db    147 GCCACTGAAAAAGGAAAAGAAAGACTCTGGTCATG--GGCCAAATGCTATTNTACACTCGGGTA 90

Qy   1373 CCACGGCGAGGAGGCTGCTGCCGCACTCAGGGCCTACTACGCAGTCACTCCCTTGACGCGCCT 1432
          || ||  | ||  || ||| | |||| ||||  ||| |||| || |  || ||| ||| |||
Db     89 CTCTGGCGGTGGATCTGCTCTAGTGACGCCATACTACACTGTGTCTCCCATTTGACGCTAT 30

Qy   1433 CAGCAAGCAAGCTCAGAGACG 1451
          |||||||  || |  ||| ||
Db     29 CAGCAAGAAGGAGATAGGCCCG 11
```

RESULT 3

```
LOCUS       CNS0766A                1021 bp    DNA     linear   GSS 07-JUL-2001
DEFINITION  CNS0766A
            T7 end of clone BB0A001C03 of library BB0AA from strain CBS 4732
            of Pichia angusta, genomic survey sequence.
ACCESSION   AL431000
VERSION     AL431000.1  GI:12214412
KEYWORDS    GSS.
SOURCE      Pichia angusta
  ORGANISM  Pichia angusta
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1  (bases 1 to 1021)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 1021)
  AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 13. Pichia
            angusta
  JOURNAL   FEBS Lett. 487 (1), 76-81 (2000)
  MEDLINE   20584723
   PUBMED   11152888
REFERENCE   3  (bases 1 to 1021)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..1021
                     /organism="Pichia angusta"
                     /strain="CBS 4732"
                     /db_xref="taxon:4905"
                     /clone="BB0AA001C03"
                     /clone_lib="BB0AA"
                     /note="end : T7"
     misc_feature    <1..>927
                     /note="similar to P07337 ( Beta-glucosidase precursor BGLS
                     )| Kluyveromyces marxianus 1"
                     /evidence=not_experimental
BASE COUNT      231 a    278 g    247 t    260 c
ORIGIN
```

Query Match   74.6%; Score 254.6; DB 17; Length 503;
Best Local Similarity 74.6%;
Matches 320; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
Qy   408 GGTATCGACTTCTGGCACACAAAGGCTCTCCCCAAGCATGGAGTCCCCTTCTCCCGCTTT 467
         |||||||||| ||||||||||| |||||| | ||| | ||||||||| | ||| | |
Db    75 GGTATCGACTTTTGGCACACCACCAAGCCTCTACACCACCATGGAATCCCTTCCCTGAGGCTA 134

Qy   468 ACAGATGGCCCCAACGGCGTAAGAGGGACCAAGTTTCTTCAATGGACGGTCCCTCGGGCCTGC 527
         | || |||||||||||||| |  | |||  |  |||||||||| || | || | || || |
Db   135 TCAGACGGCCCCAACGGCGTCCGTGGACTCCCCGGACCCGTTCTTCAACGCGTCAGGGCGGCCTGC 194

Qy   528 TTCCCCTTGGCGCACGTCGCTCGGTTCCACATTCAACCAAAACTCTGCTCGAAGAGGCAGGT 587
         ||| ||  |||||| | |||||| || | ||| || |  | ||| | ||||||||||| |||
Db   195 TTTCCCTGCGGCGGGACGGCGCTGGGAGCCCACCTTCAAACCCTCCTGGAAGAGGCCGGC 254

Qy   588 AAGATGAATGGGCAAAAGAGGCCATCGCTAAGAGTGCCCATGTGATCCTTCGGCCCGACTATC 647
         | || | ||||| |  |||| ||||||| ||| |||| || ||||||| |||||||||| ||
Db   255 CCCAAGAATGGGTGCAGGACGCCAAAGAGCGCCCATCGCCATGCCATCCTCGGCCCGACCATC 314

Qy   648 AACATGCAAGCGCTCCCCTCTCGGTGACGTGGCTTCGGACGTCGAATTGGTGAGGATCCGTTC 707
         ||||||||  || ||||||| |  |  |||||||  |   |||| | || |||||||| ||
Db   315 AACATGCAGCGTGGCCCCTCGGCGCCGTGGCTGCTCGAGTGCCGGAGGATCCTGTTG 374

Qy   708 CTGGCGCGGCGCTTGGGAGCTGCGGCTCTCATCCCGCGGCCATTCAGGACACTGGAGTGCAGGGCT 767
         |||||  ||  | ||||| | || || |||||||| |||||||||| ||| |  || |||| ||
Db   375 CTGGCCGGGACTCGGCGCCGCCCCATCGTGCGCCGGCGCATTCAGAGACGACCGGCGTCCAGGCC 434

Qy   768 ACGATCAAGCACACTTTTTGTGCAATGATCAGGAGGAGGAACAGGCGGCATTCAGAGAGCAC 827
         || |||||||| || ||| |||| | |  || ||| | || |||| ||  ||||||||||
Db   435 ACCATCAAGCACTTTGTGTGTCAACGACGAGGAGCACACCGGCCGCCCGACAGACCATC 494

Qy   828 GTCACGCGAG 836
         || ||| ||
Db   495 CTGACGGCG 503
```

ORIGIN

Query Match   54.6%; Score 227; DB 17; Length 1021;
Best Local Similarity 54.6%;
Matches 513; Conservative 3; Mismatches 414; Indels 9; Gaps 3;

```
Qy  1745 GTACGTAGACGACCAGCTCGTCGTCGACAACGGCCACCAAGCCAGGTCCCCGGGATGCCTT 1804
         ||  | |  |||||||| |  | |||| || | ||||| | | |||| || | ||||  ||
Db     3 GTTTTTGGACGACCAGTTGATCATTGACAAATTGCACAATCAGAAAACCTGGCGAGTCTTT 62

Qy  1805 CTTCGGCTCCGCCACCCGCCGAGGAGACAGGGCCCGCATCAATCTCGTCAAGGGCAACACGTA 1864
         ||| |||  | ||| ||||| |||| || ||| || |||| || |||  ||| ||||| |
Db    63 CTATGGAAAAGGAACAGTGGAGGTGCGAGGATCGTATATTGTGCAAAGCCGGAACACCGWA 122

Qy  1865 CAAGGTTCAAGATCAGAGTTTCGGCGTCCGCACCCACCTTACAGCCT---CAAGGGGGACACCAT 1921
         | ||| |||  |||  | ||||| |  |||  || | |||| |  ||    ||||  |||||
Db   123 VTGGTTGCGTATCGAGTTTGGTTCCGTTCGTCTCCCAACGAGCATGTTGCACCAAATCGAATCCTGT 182

Qy  1922 CGTCCCCGGCCACGGCTCCCTCCCGGTCGGCGGTGCCAAGGTCATTGACGAGCAGGCCGA 1981
         |  |||  | ||||| |  | | |||  | || | |  ||| || || || | ||  |
Db   183 CTTTTTTCCGGAGCTGGGGCCTGATGTTTTGGCGGTGTTGCCGGTGTTGCCGGTGTTGCCGGTGTTGCCGGT 242

Qy  1982 AATCGAAAAGTCCGTCGCCCTCGCCAAGGACGCACGAACCAGGTCATCACTCTGCCGCGGCCT 2041
         | || |  ||  ||| | |||  ||| ||| | || ||||||| ||  |||| || |||||
Db   243 GATTCAGCAGGCAGCAGAGGCAGCTCGCCAAGCTTCACAGGTTATTCTCGTTACAGGGCT 302

Qy  2042 TAACGCCGACTGGGAGACCCGAGGCGCCCGACCCGCCGAGCGGATGAAGCTCCCCGCGGCGTGCT 2101
         | || |||||  ||| | | |||  | || | | || || ||| ||||| ||||| |||||
Db   303 CAATAAGGAATAGGAATCTCGAAGGATCGACAGGAGAGAGACAATGAAAACTTCCGCGGCCATAA 362

Qy  2102 GGACCAGCTCATTGCCCACGTCGGCCCGCGGCAACCCAAAACACCGTCGTCGTCGTCATGCAGAC 2161
         || ||| ||  |  |  | |  | | |  |||  | | | |||| | | | | | |  ||| |
Db   303 CAATAAGGAATAGGAATCTCGAAGGATCGACAGGAGAGAGACAATGAAAACTTCCGCGGCCATAA 362
```

```
RESULT 4
LOCUS       CNS0772N        994 bp    DNA        linear    GSS 08-JUL-2001
DEFINITION  T3 end of clone BB0AA015H01 of Pichia angusta, genomic survey sequence.
ACCESSION   AL433353
VERSION     AL433353.1  GI:12216767
KEYWORDS    GSS.
SOURCE      Pichia angusta
  ORGANISM  Pichia angusta
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1  (bases 1 to 994)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvoglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 994)
  AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 13. Pichia
            angusta
  JOURNAL   FEBS Lett. 487 (1), 76-81 (2000)
  MEDLINE   20584723
   PUBMED   11152888
REFERENCE   3  (bases 1 to 994)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..994
                     /organism="Pichia angusta"
                     /strain="CBS 4732"
                     /db_xref="taxon:4905"
                     /clone="BB0AA015H01"
                     /clone_lib="BB0AA"
     misc_feature    <1..>994
                     /note="end : T3"
                     complement(<24..>992)
                     /note="similar to P07337 [ Beta-glucosidase precursor BGLS
                     ] [ Kluyveromyces marxianus ]"
                     /evidence=not_experimental
BASE COUNT      291 a    235 c    192 g    276 t
ORIGIN
```

```
Query Match     6.3%;   Score 187.4;  DB 17;  Length 994;
Best Local Similarity  51.2%;   Pred. No. 4.4e-16;
Matches  490;  Conservative  0;  Mismatches  461;  Indels  6;  Gaps  2;
```

```
Db    363 CGACCGCGCTAATAGACTCCGTCGTTTGACGCAAATCCGTGCACAGTCGTGTGATTCAGGC 422
          ||  || | || | || | |||||  |  |||| | ||| ||| ||| | || ||  |
Qy   2162 GGGCACCCCCGAGGAGAGATGCCTGGCCTCGCGACGCCACGCCCGCGTCATCCAGGCGGTA 2221
Db    423 AGGGACGGCCAGTTGAAATGCCTTGGGAGAGCAAAGGCTTTACTGTACACCTCTTA 482

Qy   2222 CGGGCGGCAACGAGACGGGCAACTCATTGCCGACGTCGTCGTTTGCGACTACAACCCCTC 2281
Db    483 CGGGCGGGAAACGAAACAGGCACGGGAAATTGCAGACGTCATTTTCGGCAACGTCAACCCTGC 542

Qy   2282 GGGCAAGCTGTGTCCCTCAGCTTCCCAAGCCGCCTGCAGGACAACCCCGCGTTTTCTCAACTT 2341
Db    543 TGGGAAGCTGCCCCTTACCTTCCCAAAACCGCTTAGAAGATTGTCCCTGGCTACCTAAAATTT 602

Qy   2342 CCGCACCGAGGCGCCGGCGCACGCTGTACGCCGCGAACGTCTACGTCGGCGTACAGGTACTA 2401
Db    603 CGGAACAGAAAACGGCAGGGTTCTCTATGGTCAGGATGTTTACGTTGGCGTACCCGTTTTA 662

Qy   2402 CGAGTTTGCCGACAAGGAGTCAATTCCCTTTTGGGCCACCTGTCCTACACCACTTT 2461
Db    663 CGATACACTGGGTGTCTACGCGTGTTTCCATTCGGCCGTGAGCCTACACCTCTTT 722

Qy   2462 TGCCTTTTCCAATCTCTCCGTGTCTCACACAAGGACGGCAAGCTGAGCGTGTCCCTCTCCGT 2521
Db    723 TGCTCTGTCTCGATTTAAAGGTG---CACTACGACGACATGCTCGAGGTTGAAGTTCTGGC 779

Qy   2522 GAAGAACACCGGCGTCCGTGCCCGCGCACAGGGTGGCCCCAGCCTCTACGTCAAGCCCCTCCA 2581
Db    780 GAAGAATACAGGTGAGCCAGGTGCGGTCGTTCAGGTTTACATTCATCACAACTG 839

Qy   2582 AGCGGCGCAAGATTAACCGCCCGTCAAGGAGCTCAAGGGGCTTCGCAAAGGTCGAACTGCA 2641
Db    840 CCCAGC---AATTAAGCGGCCATTCAAAGAGTTGAAAAGGATTTAGCAAAGTATTTCTGGA 896

Qy   2642 GCCCGGCGAGACGAAGGGGGGTGACAAATCGAGGAGGACAGGA 2680
Db    897 RCAGGAAGCTCCAGAAAATTGTGTATATCGATCAACGGGA 935
```

```
Qy    960 CTTGATGGAGGATGCTTCGAAAAGAAATGGGGGCCGTTGTGGCAGGCCTTGGGAGTGGCCTAATCATCATGAGCGACTGGTAC 1019
          || || ||| | || |||||| |||||  ||| |||||||||| || ||| |  | || | |||
Db    977 CTCACTGGTGTGTCGTTGCGAGAGGGAAATGGGGGTTGGAATGGTTTGGTTATGTCCGACTGGTTT 918

Qy   1020 GGCACATACAGTACCACACAGAAGCCGTTGTGGCAGGCCTTCGACCTTCGAGAT;GCCGGACCT 1079
          ||| |||| | ||| ||||| ||| ||||||| | |||||| | ||| | || | ||||| ||
Db    917 GGTGTCTACTCGATCAAAAAATTCTATTGATGCAGGTTTGGCACCTTGGACGTTCCCGCGTT 858

Qy   1080 CCACGCCTTCCGGAGGAGAAACACTCAAGTTCAACGTCTCCAA---CGGAAAGCCCTTTATC 1136
          | |||| ||| | ||| | |||| ||| || ||| |||| ||   ||| || || |||
Db    857 CCAATCATGAGAAAAATGCAGTTGTTACACGGCTATCAATGCCAGAGAGATCAGTATA 798

Qy   1137 CACGTCATTGACCAGAGGGGGCTAGGGAAGAGTTCTTCAGTTCGTCAAGAAAGTGTGGCCTCC 1196
          |||  ||||| | ||| | | | ||||| | ||| |  ||| |||  || |||| ||||||
Db    677 GACCTGATTGATGAAAGAGTTAGAAACGTGCTGAAACTTGGTGAACTTGGTCTATGCAAAGC 738

Qy   1197 GGAGTGACGGAGGAACACGGCCCCGAGACGAACTGTCAACAACACCCCGAAACGGCACGCTCTC 1256
          || |||| ||||| || | ||  | || ||||  || ||| ||| | | ||||| |||||
Db    737 GGCATTCCTGAAAATGCCTCTGAAGATTCTAAGAAATAATGTTCCTGCAGGTGTATTGCT 678

Qy   1257 CTCCGGAAGGTTGGGCAACGAGGGCATCGTCGTGCTGCTCGAAGAAGCGAGAACAACGTTCTGCCC 1316
          || | |||| || ||  ||| | ||| || | ||| || | | || | || || ||  |
Db    617 TTGAAAAGAACGGATAAGAATTGCTGTGATGTGGCCCTAACCGTAAGAGTTCCTAGAAATCTTT 558

Qy   1317 TTGACGAAGAAGAAGAAGAGACGGTGAGTGTGTCGGCCCCAAACGCCAAGCGGCACATACCAC 1376
          | || ||||  |||| ||| | || || || |||||||   || | || | ||| ||
Db    557 GGTGGGGGCTCTGCCTTCGATGAATCCTTATTATGCTACAAACAATTTTGATGGTATTGCT 498

Qy   1377 GGGGAGGCTCTGCCCGCACTCAGGGCCTACTACCGCAGTCACTCCTTTGACGGGCCTCAGC 1436
          ||  || | || ||| | | | |||| |  ||| | |  |||||  || ||||  | |
Db    497 GATAAAAATTACAGAGGCCCCCACTCTATGCTACGTGATGGTTGCTCAATTGAGAAAACCTGCTC 438

Qy   1437 AAGCACGCTGAGACGGCGCGCCCATCGTACACGTCGGCGCCTACAACGCCGTTCCTCCCATT 1496
          | | || | || | || |  |  || | | || || | | || | || || | ||
Db    497 AAGCACGCTGAGACGGCGCGCCCATCGTACACGTCGGCGCCTACAACGCCGTTCCTCCCATT 1496

Qy   1497 CTAGGCGAGCAGTGCCTCAGCGCCGAGGCGCTTGTCTACCTTCACCTTCAAACGGTTCCGAAGGGGTTTATTGCAAGGTTTACAAC 1556
          | ||| ||| | | ||| || |  | | | ||| |   |||| || ||  || |||| | | ||||  |  ||| ||  |  ||
Db    437 GATCTCGGGCACCTTGTCTACCTTCAACGGTTCCGAAGGGGTTTATTGCAAGGTTTACAAC 378
```

```
RESULT 5
CNS075SU/c
LOCUS       CNS075SU               970 bp    DNA     linear   GSS 07-JUL-2001
DEFINITION  Kluyveromyces lactis clone BA0AB038C08 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL430516
VERSION     AL430516.1  GI:12213710
KEYWORDS    GSS.
SOURCE      Kluyveromyces lactis.
  ORGANISM  Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 970)
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 11.
            Kluyveromyces lactis
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
  MEDLINE   20584721
   PUBMED   11152886
REFERENCE   3  (bases 1 to 970)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..970
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB038C08"
                     /clone_lib="BA0AB"
     misc_feature    complement(3..>968)
                     /note="similar to P07337 [ betaglucosidase ]"
                     /evidence=not_experimental
BASE COUNT      291 a    221 c    174 g    284 t
ORIGIN
```

```
Qy Match
Best Local Similarity  53.5%;          Score 184.6;  DB 17;  Length 970;
Matches 487; Conservative 0; Mismatches 404; Indels 20; Gaps 4;  Pred. No. 1.1e-15;
```

```
Qy  1557 GAGCCCCCTGGGTACCCCTTAACCCCCACCAGCACAATTGACGAGCGTCTTCTTCACCAAGACGGAC 1616
         ||  |||   ||    |||  |||      |     ||  |   |||  ||  |  |||||  ||
Db   377 CTTCCTGCTGGAATCGCGTCTGACCGTGTCATTGTGTTGATGAGTTTGCATTGAAATTCTACGAGA  318

Qy  1617 ATGCAACCTGGTGGACTACTACCACCCCCAAGGGCGGCAGACA---CGTGGTACGGCGACATG 1673
         ||     | |||  ||| |   | | || ||  | |||  |||   |  ||| | ||| ||
Db   317 TTGAGTGTTGTTTGATTAGACAAAAACGACAACAACACGTGTTTCATGATTATATC  258

Qy  1674 GAGGGCACTGACCACGCCCCGACGAGGACTGCACCTACGAGACGTCGGCCTCGCTGCGGC 1733
         ||||| | || || |  ||  |  | || || ||||||||||| | |||   | | | |
Db   257 GAGGCGGCGAGTTACCGCCAAAGAGACACTGGCACCTACGAGATAAAGCGTTAGTTGTCTTGGA  198

Qy  1734 ACGGCAAAGGCGTACGTAGACGACGCTCGTCGTCGTGTGGACAAGCGACCACCAAGCAGGTCCCC 1793
         || |||||   |  |  | || || | | ||   | |  |||| |   | |   ||| || |
Db   197 ACCGGCACAATTTTTCATTAACGGTAAATTGGAGATTGATGACAAAAATAATCAAGTCCTT  138

Qy  1794 GGCGATGCCTTCTTCGGCGTCCGGACCACCCGCCAGCACGGGGCCGCATCAATCTCGTCAAG 1853
         || |||| || |  ||| | | | || ||   | ||| || |  |||||||||| ||||||
Db   137 GGATTGCACGCCCTGGGTGCCTCTACCGGCCCCCAAAACCAAGAGAGATTTGATATGAACAAG  78

Qy  1854 GGCAACACGTACAAGTTCAAGATCGAGTTCGGGTCCGACACCCACCCTCAAG 1910
         || | |   |||||  | |||||  |||| || |||| |||||||| | |||
Db    77 GGAGAGTCTTTCAAAATACAAGATTGAGTTTGGTCGGGTCCCAACGTTCACTTTGGAG  21
```

```
Qy   583 CAGGTAAGATGATAGGGCAAAAGAGGCCATCGCTAAGAGGTGCCATGTGATCGTGGCCCGA  642
         |||||||  ||||| |||| | || ||| | ||||||| || || | ||||| ||| ||||
Db   970 CAGGTAAAATTGATGGGTAAGGAAGCCGTGGCCAAGAATGCTGCTGTGAT-TTTGGACCCA  912

Qy   643 CTATCAAGCATGCAAACCTCCCCTCTCGTGAAATCTACGGCACGTCGGATGCGGCTTCCAGATTGCCTGTG 702
         |    ||  ||||||||  | ||| |  || ||  |  |  | | | |||| ||||| ||||  |
Db   911 CGGCTAACATGCAAACGTGCGTCCTTTGGGGTGGTCGTGGCTTTGAATCATTTTCTGAAGATC  852

Qy   703 CGTTCCTGGCGGGGCTTGGGAGCTGCGCGCTCCTCATCCGGCGGCATTCAGAGC-ACTGGAGTG 761
         | |   | |  | || |   | ||   || || | |  ||  ||||| |||||| || | ||
Db   851 CATATTTAGCTGGTGTTGCTACTGCTTCGGTTGTTCAAGGAATCCAAAATCTGAGGGTATT  792

Qy   762 CAGGCTACGGATCAAGCACTACTTTTTGTGCAAATGAATCAGGAGGACGCCATGATGTGACAG 821
         |||||||||| ||||||||   |  |||   ||||| ||| | ||| | | ||||  ||| |
Db   791 GCTGCTACGGTAAAGCACTTTGTCTTGTATTGATCTAGAAGATCAACGTTTCTCTTCGGAAT  732

Qy   822 AGCAATCGTCACGGAGCGGGGCTCTCCGTGAAAATCTACGGCCACTCCGTTCCAGATGCTGTG 881
         | | |  | | | ||| ||   ||| | ||||| |  ||   | || ||||||||| | ||
Db   731 GCCACGCTCTCTGAGAGAGCGTTGCGAGAAATATACTTGAGACCTTTCAGATTTGGGCAGTC  672

Qy   882 CGAGACTCCCACGCGGGTGCGTTCATGACGGCGTACAAATGGCATCAATGGCGTGTCGTGC 941
         | |||  || |  | | || || || |  ||   ||||| ||||| ||| | | | |||
Db   671 AAGAATGCTGAATCCTGTTTGCTTGTTTAATGTCCGCTTAACAATAAGGTGAACGGCGAACATTGT 612

Qy   942 AGCGGAGAACCCTAAATATCTTGATGGGAGATGCTTCGAAAAGGAATGGGGTTGGATGGGCCTA 1001
         | | |||| |   ||| | ||| || |  |  | ||| |||||| ||| | ||||  || | |
Db   611 TCTCAGAACAAAAAAATTGTTATTGGACATTCTTAGAAAAAGAAATGGAATTGGGAATGGTATG  552

Qy  1002 ATCATGAGGCGAGCTGGTTACGGCACATACACTACCACAGAAGGCCGTTGTCGCAGCCGCTCGAC 1061
         || | | || | | || || | | || |   || || | | || ||||| || |||| | |
Db   551 ATTAATGTCTGATTGGTTTGGTACTTTACCACACTCGTGCTTCTATCAAGAGAATGGTCTGGAT  492

Qy  1062 CTCGAGATGCCCGGACCTCCAGCGCTTCCGGAGGAGAAAACACTCAAGTTCAAGTC------ 1115
         || || | || |||  | || |  |||| |   || | |||| |||||| |||
Db   491 ATTGAGTTCCCAGGTCCAACGAGATGGAGAACCTGTTTGTTCACATTCTCTAAAT  432

Qy  1116 TCCAACGGGAAAGCCCTTTATCCACGTCATTGACCAGAGAGGGCTAGGGAAGTTCTTCAGTTC 1175
         ||||  || |||   || |  | | ||  || |||   |  |  |||  |||| | | |
Db   431 TCCAGGGAACAAATTTCTATTTATGACGTTGACGACCGGTGTTAGACAACGTTTTGAAAATG  372

Qy  1176 GTCAAGAAGTGTGCTGCCTCCGGAGTGACG---------GAGAACGGCCCCGAGACG 1223
         | | || | | || |  | | | || || |         |  ||| | ||||||| |
Db   371 ATTAAGTTTGTTCTTGACAACCAGAAAAAGACGGGGGGATTGTTCAAAATGGTCCAGAAACC  312

Qy  1224 ACTGTCAACACACACCCCCGAAACGGCAGCGTCTCCCCCGGGAAGGTTGGGCAACGAGGGCATG 1283
         ||  |||| | |||| ||||||||| | || ||| ||| ||||| || || ||| | || |||
Db   311 ACATCTAACAAATAGTAAAGAGACCTCTGAATTGCTTAGAAGATAGCTCCGGGATTCCATT  252

Qy  1284 GTGCTGCTGAAGGAAGGAGAACAACGTTTCTGCCCTTGAGCAAGAAGAAGAGAGACGGGTGATT 1343
         || |||||||| ||| |||||||| | ||||| || | || |||| || || || | | |||
Db   251 GTCTTGTTGAAGAACGAGAAATTCGATCTTCTACCTTTGAAAAAAGAGAATCAGTCGTCGTT  192

Qy  1344 GTCGGCCCCAACGCCAAGCCACATACCACCGCGGAGCCTTCGCCGACTCAGGGCC 1403
         || ||  ||||| | | ||  |   | |||| || | || || |
```

```
Db    191  ATTGGTACCAAAACGCCAAAGCAAAAGGCTTCTTCTGGAGGTGGTTCTGCATCTATGAACTCT  132
Qy   1404  TACTACGGCAGTGCACTCCCTTTGACGGGCCTCAGCAAGCAGCAGCAGCGCGCCCATCGTAC  1463
            |||||  |||||||||||||||||||||||||  |  |||||||||||||||
Db    131  TACTATGTTATTTCACCTTACGAAGGGATCGTCAAGAAAGTGGGCAAGGAAGTACCATAT   72
Qy   1464  ACCGGTCGGCGC  1474
            ||| |  |||
Db     71  ACAAATTGGGGC   61
```

FEATURES   Location/Qualifiers
     source   1..1045
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB017E12"
                     /clone_lib="BA0AB"
BASE COUNT   314 a   243 c   182 g   300 t   6 others
ORIGIN

Query Match   5.9%; Score 176.8; DB 17; Length 1045;
Best Local Similarity   51.2%; Pred. No. 1.1e-14;

```
Matches  410;  Conservative   1;  Mismatches  390;  Indels   0;  Gaps   0;

Qy   1959  AAGGTCATTGACGACCAGGCCGAAATGAAATCGTCGGTCGCCCTTCGCCAAGGAGCACGAC  2018
            |||| |  |||| ||||||||||||| ||||||   ||||||||||  ||||| ||||||
Db    825  GATATGACTTTGCCAGGAAGAAACAAAACGGATCCTGTCCGTGCTGTTTTGGAGGCCAATCCT  766

Qy   2019  CAGGTCATCATCTGCGCGGGGCCTTAACGCCGGACTGGGAGACCGAGGGCGCCGACGCGTC  2078
            |||| ||||| ||||||||  ||| |||||||||||| || |||| ||| ||||| |||||
Db    885  AAAGCCAATTTATGTAATTGGTTTGAACGCAGAAATGGGAAATCTGAAGGTCATGATAGGGAA  826

Qy   2079  AGCATGAAGCTCCCCGCGCGTGCTGGACCAGGCTCATTGCCACGGTCGCCCGGCGGAACCCA  2138
            |||  |||||||| ||||||  ||| ||| ||||  ||||||||| ||||| ||||||||
Db    945  AAGCCTATTGATGTTGGAGCGCGCCCCTCGCCGGGTGGCTCGACGGCGGCCACG  886

Qy   2139  AACACCGGTCGTCATGCACGACGGGCGCACCCCGAGGAGAGATGCCCTGGCGTCGACGGCCACG  2198
            |||||| ||||||||||||||||||||||||  |  ||||||||| |||||  ||||||||||
Db    765  AACACTGTTAAATTGTCAACCACTGTCTGTACTCCTGTCGAGTTCCCATGATTACAAAAAGTT  706

Qy   2199  CCCGGCCGTTCATCCAGGCGCTGGGTTACGGCGCGGAACGGAGCACGAGGCAACTCCATTGCC  2258
            |||||||||||||||||||||||||| |||| ||||||||||| ||||  |||||||||||
Db    885  AAWHHTCTCATTCAAGCTTGGTTACGGACGGGTAATGAGGTHGGGCAAATGCTATTGCTGATGTC  826

Qy   2259  GTCTTTGGCGACTACAACACCCCTCGGGCAAGCTGTCCCTCAGCCTTCCCCAAGCCGCCTGCAG  2318
            |||||||||||||||||  ||||||||||  ||||| |||| ||||||   ||||||  ||||
Db    645  TTGTATGGTGACGTTAATTCCAAGTGGGTAAATTATCTCTCAATGGCCAATAAAACCTTGAA  586

Qy   2319  GACAACCCGCGGTTTCTCAACTTCCGCACCGAGGCGCCGGGCGCCACCTGTACGGCGGAGGAC  2378
            ||| |||||||||| |||||||||||||| ||||||||||||||||||||| |||||||||||
Db    585  GATAACCCAACTTATTTTGAACTTCAAAGACTGAAATTTGGCAGAGTTCTCTACGGTGAGGAT  526

Qy   2379  GTCTACGTCGGGTACAGGGTACTACGGAGTTGCCGACAAGGAGCGTCAAATTTCCCCTTTGGC  2438
            |||||| ||||||||||||||||||||||||||||||| ||||  |||||||||||||||| |
Db    525  ATTTTCATTCGGATACAGAGATTTCTATGAGAAGAAGAGTTGCAAAGAGTTGCATTNCCCTTCGGT  466

Qy   2439  CACGGCCCTGTCCTACACCACTTTTGCCTTTCCAAATCTCTCCGTGTCTCACAAGGACGGC  2498
            |||||||    | || ||||||||||||||| | ||||||||||||||| |||||||||||
Db    465  TACGGGTTATCTTATACTGAAATTGCTTTAATCTACACAGTACAAACGGATAGGAGGAG  406

Qy   2499  AAGCTGAGCGTGTGTCCCTCTCCGTGAAGAACACCGGCTCCGTGCCGCCGCACAGGGTGGCC  2558
            |||||||||| ||||||||||| || ||||||||||||  |||||| || |||||||||||
Db    405  GTTAATCTCGTTCGTTTCTGTTAAATGTAAAGAAACCGCAGGTGAGGCATGGCATGATCCGAAGTC  346

Qy   2559  CAGGTCTACGTCAAGCCCCCTCCAAGCGGCCAAGATTAAACCCGCCCCCGCCCCTTTGCAG  2618
            |  || || |  | ||||||||  |||||||||| |||| ||||||||||||||||||||||
Db    345  GTACAAAATTTACATTGCTCGCGACTGAATCATCTTTTTCTAGGGCGAGTCAGTACAAGAGTTGAAG  286

Qy   2619  GGCTTCGCAAAGGTCGAAACTGCAGCCCGGCGACGAAGGCGGGTGACAATCGGAGGAGCAG  2678
            |   ||   ||||||||||  |||| |||||| ||| | ||| |||| |||| |||||||
Db    285  GGATTCAAGAAAAATTTGCTTCAAACCTGGCAAAACGAAACAGTCAAAGTTGATTTAATC  226

Qy   2679  GAGAAGTACGTCGCTGCGTGGGTATTTTGATGAGGAGGAGGGGATCAGTGGGTGTGTCGAAAAGGGT  2738
            |||||||  || |  || | || |||| || | || |||| || |||| || ||||||  |||||
Db    225  TTGAAGGATTCTATTTCCTTCTTCCGATGAAGAAGAAGTGGTAAGTGGTTCAGAAGCTGGT  166

Qy   2739  GACTATGAGGTTATCGTGAGC  2759
            ||    ||| |  || |  ||
Db    165  CAGTACAAAAGTCTTGGTAGGC  145
```

RESULT 7
CNS073UC/c
LOCUS   CNS073UC   952 bp   DNA   linear   GSS 07-JUL-2001
DEFINITION Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL427978
VERSION   AL427978.1 GI:12211172
KEYWORDS   GSS.
SOURCE   Kluyveromyces lactis.
  ORGANISM Kluyveromyces lactis.
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1 (bases 1 to 952)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,

---

RESULT 6
CNS07387/c
LOCUS   CNS07387   1045 bp   DNA   linear   GSS 07-JUL-2001
DEFINITION Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL427181
VERSION   AL427181.1 GI:12210375
KEYWORDS   GSS.
SOURCE   Kluyveromyces lactis.
  ORGANISM Kluyveromyces lactis.
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1 (bases 1 to 1045)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,P., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvrglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2 (bases 1 to 1045)
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 11.
            Kluyveromyces lactis
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
   MEDLINE   20584721
   PUBMED   11152886
REFERENCE   3 (bases 1 to 1045)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES   Location/Qualifiers
     source   1..1045
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB017E12"
                     /clone_lib="BA0AB"

DEFINITION clone BA0AB017E12 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.

DEFINITION clone BA0AB023A01 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.

```
          AUTHORS   Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
                    de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
                    Malpertuy,A., Neuvegise,C., Ozier-Kalogeropoulos,O., Potter,S.,
                    Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
                    Wincker,P. and Weissenbach,J.
          TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
                    yeast species for molecular evolution studies
          JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
          MEDLINE   20584711
          PUBMED    11152876
          REFERENCE 2  (bases 1 to 952)
          AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
                    Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
                    Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
          TITLE     Genomic exploration of the hemiascomycetous yeasts: 11.
                    Kluyveromyces lactis
          JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
          MEDLINE   20584721
          PUBMED    11152886
          REFERENCE 3  (bases 1 to 952)
          AUTHORS   Genoscope.
          TITLE     Direct Submission
          JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
                    2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
                    seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
          COMMENT   This GSS is part of a random genomic sequencing program of thirteen
                    yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
                    exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
                    Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
                    lactis, Kluyveromyces marxianus var. marxianus, Pichia
                    angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
                    Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
                    5 kb were prepared and both extremities were sequenced. See
                    keywords for description of this sequence and for the sequence of
                    the other extremity of this insert.
          FEATURES             Location/Qualifiers
               source          1..952
                               /organism="Kluyveromyces lactis"
                               /strain="CLIB 210"
                               /variety="lactis"
                               /db_xref="taxon:28985"
                               /clone="BA0AB023A01"
                               /clone_lib="BA0AB"
          BASE COUNT    295 a    217 c    167 g    267 t      6 others
          ORIGIN

          Query Match      5.9%;  Score 175.4;  DB 17;  Length 952;
          Best Local Similarity  52.0%;
          Matches  386;  Conservative  3;  Mismatches 354;  Indels   0;  Gaps   0;
                                          Pred. No. 1.8e-14;

Qy   2013 CACGACCAGGTCATCATCTCGCGCGCGGGCCTTAACCCCGACTGGGAGAGCCACCGAGGGCCGAC 2072
              |  |||||||        |      ||||        |||        ||
Db    940 CATGATAAAGCCATTTATGCATCGGTTTGAAATGCTGAATGGGAATCTGAGGGCCATGAT  881

Qy   2073 CGCGGGAGCAGTGAAGCTCCCCGGCGGTGGTGCTGCTGGCACAGCTCATTGCCGACGTGGCCGCCGCGCGCG 2132
           |  ||       ||   ||        |||    ||||   || |||     ||
Db    880 AGGGAAGATATGACGTTGCCCAGCAAGACAACGATTCTAGTTCGTGCTGCTGTAGGAGGCT  821

Qy   2133 AACCCAAACACCGTCGTCGTCATGCAGAGCGGGACACCCCGCGAGGAGATGCCCTGGCCTCGAC 2192
           ||||| ||||| || | || |   |||| |||  |  |   |   ||| |||  |||  |||
Db    820 AACCCTAACACTGTAATTGTCAACCAGTCTGGGTACTACTCCTGTCGTGAGTTCCCATGGTTACAA  761

Qy   2193 GCCACGCCGCCGCCGTCATCCAGGCCTTGGTACGGGCGCCAACGAGACGGGCGAACTTCCATTGCC 2252
            || ||   |||| ||||   | |||||| ||| ||   |||| |||| ||| ||| |||||||||
Db    760 AAAGCTAAAGCTCTGGTTCAAGCTTGTTGTACGGACGGCAATGAGCTGGGTAAGTGCTTATTGCCT  701

Qy   2253 GACGTCGTCGTTCTTGGCGACTACAACCCCTCGGGCAAGCTCGTCCCCTCAGCGTTCCCCAAGCGC 2312
           ||||| |   ||| ||| ||   | |   | |   |||                |||     ||
Db    700 GACGTCTTGTATGGTGATNGTGTYCAAAATGGGAAAATNTATNCTCCTCTCAYGGCCATTAAAG  641

Qy   2313 CTGCAGGACACACCCCGCGGTTTCTCAACTTCCGCACCGAGGCGGGGGCGGCGGGGCTGTACGGC 2372
           || ||| ||||| ||  | |   |  |||||| | |  |||| | |||  ||     |||||||
Db    640 CTTGAAGACAAATCCAGCTTATTTGAACTTCAAGACGCGAATTTGGCAGAGTTGTGTCTACGGT  581
```

```
          RESULT 8
          CNS072ML
          LOCUS       CNS072ML                 812 bp    DNA     linear   GSS 07-JUL-2001
          DEFINITION  Clone BA0AB012F11 of library BA0AB from strain CLIB 210 of
                      Kluyveromyces lactis, genomic survey sequence.
          ACCESSION   AL426403
          VERSION     AL426403.1  GI:12209597
          KEYWORDS    GSS.
          SOURCE      Kluyveromyces lactis.
          ORGANISM    Kluyveromyces lactis
                      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                      Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
          REFERENCE   1  (bases 1 to 812)
          AUTHORS     Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
                      Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
                      de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
                      Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potter,S.,
                      Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
                      Wincker,P. and Weissenbach,J.
          TITLE       Genomic exploration of the hemiascomycetous yeasts: 1. A set of
                      yeast species for molecular evolution studies
          JOURNAL     FEBS Lett. 487 (1), 3-12 (2000)
          MEDLINE     20584711
          PUBMED      11152876
          REFERENCE   2  (bases 1 to 812)
          AUTHORS     Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
                      Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
                      Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
          TITLE       Genomic exploration of the hemiascomycetous yeasts: 11.
                      Kluyveromyces lactis
          JOURNAL     FEBS Lett. 487 (1), 66-70 (2000)
          MEDLINE     20584721
          PUBMED      11152886
          REFERENCE   3  (bases 1 to 812)
          AUTHORS     Genoscope.
          TITLE       Direct Submission
          JOURNAL     Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
                      2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
                      seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
          COMMENT     This GSS is part of a random genomic sequencing program of thirteen
                      yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
                      exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
                      Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces

Qy   2373 GAGGACAGTCTACGTGCGGGTGGTACAGGTACTACGGAGTTTGCGACAAGGAGGACGTCAATTTCCCC 2432
           |||||  |  |   |  |||| ||| ||| | | | |||||||      ||          ||  ||||
Db    580 GAGGATATTYTCATTGGATACAGATTCTATGAAAGATAAAGCTCCAAAAAGAGTGGCATTCCCC  521

Qy   2433 TTTGGCCACGGCCTGTCCTCACACCACTTTTGCCTTTTCCAATCTCTCCGTGTCTCACAAG 2492
           |||||  ||||| | |||   || | |   |||||  |||  ||  |  | |     ||
Db    580 TTCGGTTACGGGTTATCTTATACTATACCGAAGATTTGCCTTATCGTCAAGTTCAAATAAT  461

Qy   2493 GACGGCAAGCGTGAGGGTGTTCCCTCYCCGTGAAGAACACCCGGCTCCGTGCCCGGCGCACAG 2552
           ||||| |||  | |||   || |  ||   || || ||||  ||  || || |  | | ||
Db    460 GACGAAGTTATCTCGTGTTCTTCGTTGATGTTAAGAACACAGGTGAGAAAATATGCGGGATCT  401

Qy   2553 GTGGCCCAGCGTCTACGTGCATCAAGCCACCCTCCAAGCGGGCCAAGATTAACCGCCGCCAAGGAG 2612
            |  |  |||| | |  | || ||| | |||||||| |||||||  |||| ||  |  | |||||||
Db    400 GAAGTCTACAACAAGTTTATATKGCTGCTACTGAATCATCTGTTTCTAGGCGGCAGTTGAAGGAG  341

Qy   2613 CTGAAGGGCTTCGCCAAAAGGTCGAACTGCACCCCCGGCGAGACGAAGGCGGTGACAAATCGAG 2672
           ||||||||  || | |||| | | ||    |   |  ||||||| || |  |  |||| |||
Db    340 TTGAAGGGATTCAAGAAAGTTTTACTTCAACCCGGCAAAACGAACAGCAAAGATTGAT  281

Qy   2673 GACGCAGGAGAAGTACGTGCGCTCGCTATTTTGATGAGGAGGGATCAGTCGGTGTTGTCGGAA 2732
            || || ||| |   |||| | | ||  |||| |  |||||   ||||| |  | |||| |||||
Db    280 TTAGTCTTGAAGGACTCGTGTCTCCTTCGTCAGATGAGAAGTCGGGTAAGTGCTGTTGTTCAGAA  221

Qy   2733 AAGGGTGACTATGAGGTTTATCGT 2755
            |||| ||||||| | |  ||  ||
Db    220 GCTGGTCAGTCACAAAGTCTTGGT  198
```

FEATURES
        source   1..812
                        /organism="Kluyveromyces lactis"
                        /strain="CLIB 210"
                        /variety="lactis"
                        /db_xref="taxon:lactis"
                        /clone="BA0B012F11"
                        /clone_lib="BA0AB"
        misc_feature   <92..>811
                        /note="similar to P07337 [beta glucosidase]"
                        /evidence=not_experimental
BASE COUNT   219 a   148 c   193 g   232 t   20 others
ORIGIN   Location/Qualifiers

Query Match   5.8%;   Score 171.8;   DB 17;   Length 812;
Best Local Similarity   52.0%;   Pred. No. 5.7e-14;
Matches 381;   Conservative 12;   Mismatches 331;   Indels 8;   Gaps 1;

Qy   1802 CTTCTTCGGCTCCGCCACCCGGAGGAGACGGGCCCGGCCATCAATCTCGTCAAGGGCAACAC 1861
          || | |||| |||| || || | ||| ||||| | | |||| |||||||| | ||||
Db   6 CTGTTTCGGTTGCTGGTTACCAAGDGAAAAGACTAAGACTGTTACTCTACACAAAAGGTAAATC 65

Qy   1862 GTACAAGTTCAAGATCGAGTTCGGGCTCCGGCCACCCTACACCCTCAAGGGCGACACCAT 1921
          |||| || || || || ||||||| || | || ||| ||| ||||| | || ||||
Db   66 GTATAAAGTCAGAATGNAWDDCTCTGGCCCAACTAGCC-------AACTGGTAAG 117

Qy   1922 CGTCCCGGGCCACGGCGGCTCCTCCGCGGTCGGCGGTCATTGACGACCAGGCCGA 1981
          ||| ||||| | | |||| ||| |||||| || |||||||| ||||
Db   118 TGAATTCGGAGCTCGTGCTCTCCAGGTTGGTTGCMAAGGCTTATTGATGAGATGAGGA 177

Qy   1982 AATCGAAAAGTCCGTCGTCGCCCTCGCCAAGAGAGCACGACCAGGTCATCAANCTGCCGGGGCCCT 2041
          || | || || ||| || || ||| ||| || || || ||| |||| | || || |||||||
Db   178 AATCAAGAAGGCAGCCAAAACTTGCTGCCGGGGTCACGATAAAAGCCATTTTATGTATTGGTTT 237

Qy   2042 TAACGCCGACTGGGGAGGCCGGAGGGGCCCGACCCCGGAGCATGAAGCTCCCGGCGGTGCT 2101
          |||||: ||| || ||| || ||| || | | ||| || || |||| ||||| || ||
Db   238 GAACGCCMGAANGGGAATCGAAGGTCATGATAGGGAGATATGCACTTTGCCAGGMAGAAC 297

Qy   2102 GGACCAGCTCATTGCCGACGGTGGCCGCCGCCGGAACCAAACACGGTCGTCGTCATGCAGAC 2161
          || ||| | || ||| | | ||| | || || | |||||| | || || |||| ||
Db   298 AAACGATCTGGTCCGTGCYGTTTTGGGAGGCCAATCCTRACACACTGTMATTGGTCAACCAGTC 357

Qy   2162 GGGGCACCCCGAGGAGATGGCCCCTGGCTCGACCGCCACGGCGTCATTCCAGGCCTGGGTA 2221
          |||||| | || || || || ||||:: ||| | || || |||| || || | ||||
Db   358 TGGTTACTTCCTGTGTGGCAGTTCCATGATTACCAAAMGCTTAAAGCTTCTCATTCMAGCTTGGTA 417

Qy   2222 CGGCGGCCAACGAGACGGGCCAACTCCATTGCCCGACGTCGTCTTCTTTGGGGACTACACACCCTC 2281
          ||| | ||||| || | ||| | | | | | ||| |||| |||| ||| || |||:|:
Db   418 CGGAGGGTRAATGAGTGGGCMATGCTATTGCTCGATGTCTTGTATGGTGACGTTATTCCMAG 477

Qy   2282 GGGCAAGCTGTCCCTCAGCTTCCCCAAGCGCCTGCAGGACAACCCCGGCGTTTCTCAACTT 2341
          || ||| |||| | || ||| | || || || |||||| | ||||||| || ||||||
Db   478 TGGTAAATYATCTCTCTCACATGGCCAATMAACCTTGAAGATAACCCAACTTATTTGAAACTT 537

Qy   2342 CCGCACCAGGCGCCGGCGCCGCACGCGTTACGGACGAGCGTCTACGTCGGGGTACAGGTACTA 2401
          || || | | | | || || | || | || | ||| |||| || || | ||| |||
Db   538 CAAGACTGAATTTGGCAGAGTTCTCTACGGTGCAGGATATATTTCMTTGGATACAGATTCTA 597

Qy   2402 CGAGTTGCCGACAAGGACGTCAAATTTCCCCCGTTTGGCCACGGCCTGTCTCTACACCACTTT 2461
          ||||| ||||| | || ||| |||| | || |||| ||||||| ||| || || || |||
Db   598 TGAGAAGTTGCCAAGMAGAGTTGCMTTCCCCTTCGGTTACGGGTTATCTTAAGTCTTATACTGAATT 657

Qy   2462 TGCCTTTTCCAANTCTCTCCGTGTCTCACAAGGACGGCGTCCCCTCTCCGT 2521
          ||| || || || ||| || || ||||||| || ||| ||||||| |||
Db   658 TGCTTTATCTRAATCTRACACAAGTNACAAACGGATGAGGAGGTTATTCTCTGTTTCTGTTAAGT 717

Qy   2522 GAAGAACACACGG 2533
          ||| || | ||
Db   718 AAAGAACGCAGG 729

REFERENCE
        AUTHORS

Db   718 AAAGAACGCAGG 729

LOCUS   CNS070RO   917 bp   DNA   linear   GSS 06-JUL-2001
DEFINITION CNS070RO/c
            T3 end of clone Az0AA011G05 of library Az0AA from strain CBS 712 of.
ACCESSION   AL423994
VERSION   AL423994.1   GI:12207188
KEYWORDS   GSS.
SOURCE   Kluyveromyces marxianus.
  ORGANISM Kluyveromyces marxianus
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1 (bases 1 to 917)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2 (bases 1 to 917)
  AUTHORS   Llorente,B., Malpertuy,A., Blandin,G., Artiguenave,F., Wincker,P.
            and Dujon,B.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 12.
            Kluyveromyces marxianus var. marxianus
  JOURNAL   FEBS Lett. 487 (1), 71-75 (2000)
  MEDLINE   20584722
   PUBMED   11152887
REFERENCE   3 (bases 1 to 917)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES   Location/Qualifiers
        source   1..917
                        /organism="Kluyveromyces marxianus"
                        /strain="CBS 712"
                        /variety="marxianus"
                        /db_xref="taxon:4911"
                        /clone="Az0AA011G05"
                        /clone_lib="Az0AA"
                        /note="end : T3"
        misc_feature   complement(<5..>916)
                        /note="similar to P07337 [beta-glucosidase precursor (EC
                        3.2.1.21)]"
                        /evidence=not_experimental
BASE COUNT   262 a   203 c   177 g   275 t
ORIGIN   Location/Qualifiers

Query Match   5.8%;   Score 167.2;   DB 17;   Length 917;
Best Local Similarity   52.0%;   Pred. No. 2.2e-13;
Matches 426;   Conservative 0;   Mismatches 368;   Indels 18;   Gaps 2;

Qy   684 GAGTCGATTGGTGGAGGATCCGTTCCTGGCCGGGGGCTTGGGAGGCTGGGCGCTCTCATCCGGCGGC 743

RESULT 10

```
LOCUS       CNS072H1
            CNS072H1      745 bp    DNA     linear   GSS 07-JUL-2001
DEFINITION  Kluyveromyces lactis, genomic survey sequence.
            clone BA0AB011D05 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL426203
VERSION     AL426203.1  GI:1209397
KEYWORDS    GSS.
SOURCE      Kluyveromyces lactis
  ORGANISM  Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
```

```
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   MEDLINE  20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 745)
  AUTHORS   Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
```

```
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 11.
            Kluyveromyces lactis
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
   MEDLINE  20584721
   PUBMED   11152886
REFERENCE   3  (bases 1 to 745)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..745
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB011D05"
                     /clone_lib="BA0AB"
BASE COUNT      212 a    134 c    180 g    216 t      3 others
ORIGIN
```

Query Match 53.0%; Score 165.4; DB 17; Length 745;
Best Local Similarity 5.6%;
Matches 378; Conservative 2; Mismatches 324; Indels 9; Gaps 1;
Pred. No. 4.2e-13;

```
Db  916 GAATCAATTCTCTGAAGATCCATATCTTGCTGGTATGGCTACTTCTTCTTGTTGTTAAAGGT 857
            |||                        |||                   |||

Qy  744 ATTCAGAGCAACTGGAGGTGCAGCGCTACGATCAAGCACTTTTTGTTGCAATGATCAGGAGGAC 803
            |||| ||||||  ||                |||
Db  856 ATGCAAGGCGAAGGTATTGCTGCTACCGTTAAGCATTTGTTTGTTGTAACGACTTGGAAGAC 797

Qy  804 AGGGCGCATGATGGTGCAGAGCATCGACGGAGCGGGCCTCCGTGAAATCTACGCACTC 863
                 |||                |||                  |||
Db  796 CAACGTTTTCTTCGGACTCAAATTGTTCTGTCTCAGATGTTCTCAGAGAAATTTACTTGGAG 737

Qy  864 CCGTTCCAGAATTGCTGTGTCGCGAGACTCCCAGCCGGTGCGGTTGGCCGTACAATGGC 923
              ||||                   |||               |||
Db  736 CCCTTCAGATTGGGCAGTTAAAACATGCCAAATCCTGTGTTTGCTATAATGACTGCTTTATAACAAG 677

Qy  924 ATCAATGGCGTGTTCGTCGTGCAGCCGAGAACCCTAAATATATCTTGATGGGACCCTTCGAAAGGAA 983
            ||||||               |||               |||
Db  676 GTCAATGGCGGAACATTGCTCTCCCAATCCAAAGAAGCTATTGATCGACATTTTGAGGACGAG 617

Qy  984 TGGGGTTGGGGATGGCCTPAATCATTGAGCCGACTCGGTTACGGCCACATACGTACCCAGAAGCC 1043
              |||                  |||                  |||
Db  616 TGGAAATGGGACGCGTATGTTAATGTCCGACTGGTTCGGTACATATACGGACTCCCGCAGCT 557

Qy 1044 GTTGTTGGCAGGCGTTCGACCTCGAGATGCCCACGCTTCCGAGGAGAAACACTC 1103
            ||||                   |||                 |||
Db  556 ATCAAGAATGGGTTGGATTATCGAGTTCCCTGGACCACAAGATTGGAGAGAACACGTGCTTTA 497

Qy 1104 AAGTTCAAC------GTCTCCAACGGAAAGCCTTTATCCAGTCATCATTGACCAGAGGGCT 1157
                               |||                 |||
Db  496 GTGTCTCACTCTCTCAACTCCAGAGAACAAATCTACTACTGAAGANTGTTGATGATCGGTGTT 437

Qy 1158 AGGGAAGTTCTTCAGTTCGTCTCAAGAAGTGTGCTCCTCCGGGACTGA------------CG 1205
            |||                  |||                 |||
Db  436 AGACAAGGTGCTAAAAATGATTAAGTTCGTTGTGTTGTGACAAITTAGAGAAAACAGGTATTGTG 377

Qy 1206 GAGAACGGGCCCGAGACGACTGTCAACAACGCCAAGCCGAGCTCTCCTCCCGGAAG 1265
            ||||                   |||                |||
Db  376 GAGATAGGCCCCAGAATACTACTTCAAACAACACCAAGGAAAACCTCGGACCTGTTGAGAAAG 317

Qy 1266 GTTGGCAACGAGAGGGCCATCGTCTGCTGCCTGAGAGAGAGAGACAACCTTTCTCGCCCTTGAGCAAG 1325
              |||                 |||                 |||
Db  316 ATTTGCTGCTGCACTCTATTTGTTTTATTGAAGAAACAAATATCTTACCTCTAAAGAAA 257

Qy 1326 AAGAAGAAGAGAGCGTGATTGTCGGCCCCAAGCGCCACTACCACGCGCGGAGGC 1385
            ||||                   |||                |||
Db  256 GAAGACAATATCATTTGTCATTGGCCCCAAATGCTAAAGACTAGTTCCGGTGGTGGT 197

Qy 1386 TCTGCGCACTCAGGGCCTACTACGCCAGTCACTCCCTTTGACGGCCTCAGCAAGCAGCTC 1445
              |||                 |||                 |||
Db  196 TCAGCCATCTTATGAACTCCTACTACTAAGTTGTTGTTCTCCGTATGAAGGTATCGTCAATTAAGCTG 137

Qy 1446 GAGACGCCGCCCATCGTTACACCGTCGGGCGGCCCTA 1477
              |||                 |||
Db  136 GGCAAAGAGGTCGATTACACCGTAGGCGGCCCTA 105

Qy 1821 CGGGAGGAGGACGGCGCGCGCCATCAAATCTCGTCAAGGGCAACACACGTACAAGTTCAAGATGGAG 1880
              |||                 |||
Db   37 CTCGAAAAGACTAAGACTCGTTACTCTACAAAAAGGGTAAAATCGTATAAAACTCAGAAATCGAA 96

Qy 1881 TTCGGCTCCGACACCCACCCTCAAGGGGGACAACCAATCGTCCCCGGGCCACGGGCTCC 1940
              |||                 |||
Db   97 TATGGCTCTGGCCCAACTAGCCAACTGGTAAGTGA------ATTCGGGAGACTTGGGTGCT 147

Qy 1941 CTCCGCCGTGCGGCGCTGCAAGGTCATTGACGACCAGGCCGAAATCGAAAAGTCCGTCGCC 2000

Qy 2001 CTCGCCAAGGGCAACGACCAGGTCATCAATCTGCGGCGGCCTTAAACGCCGGACCTGGGAGACC 2060
              |||                 |||
Db  148 CTCCAGGTTGGGTGTTGCAAAGGGTATTGATGCAGATGAGGAAATCAAGAGAGGCAGCAAAA 207

Qy 2061 GAGGGCGCCGGACCCGGCGGAGCAGTGAAAGCTCCCCGGCGTGCCTGGACCGGCTCATTGCCGAC 2120
              |||                 |||
Db  208 CTTGCTGCGGGTCACGATPAHAGCCATTTTATGTATTGGTTTGAACGGCAGAAATGGGAACTG 267

Qy 2121 GTGGCGCGGCCCGGACACCGAACCAAACACCGTCGTCGTCACATGCACGGACGGGCCACCCCGGACGGGATG 2180
              |||                 |||
Db  268 GAAGGTCATGATAGGGAAGGAAGATATGACTTTGGCCAGGAAGGAACAAACGAATCGTCGTCCGGTGCT 327

Qy 2181 CCCTGGCGTCGAGACCCGCCACCCCGGCGGTCATCCAGCCGCTGGTACGGCGGGCAACGAACGGGC 2240
              |||                 |||
Db  328 GTTTTGGAGGCCAAATCCTAACACTGTAATTGGTYAAACACTGGTTAACGATTCGGTTACCGGGTGGTTC 387
```

REFERENCE   AUTHORS   Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.

```
RESULT 11
LOCUS       CNS078JO      993 bp    DNA    linear   GSS 08-JUL-2001
DEFINITION  T3 end of clone BB0AA021B09 of library BB0AA from strain CBS 4732
            of Pichia angusta, genomic survey sequence.
ACCESSION   AL434074
VERSION     AL434074.1  GI:12217488
KEYWORDS    GSS.
SOURCE      Pichia angusta
  ORGANISM  Pichia angusta
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1  (bases 1 to 993)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvoglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   MEDLINE  20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 993)
  AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
  TITLE     Genomic exploration of the hemiascomyceetous yeasts: 13. Pichia
            angusta
  JOURNAL   FEBS Lett. 487 (1), 76-81 (2000)
   MEDLINE  20584723
   PUBMED   11152888
REFERENCE   3  (bases 1 to 993)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See.
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..993
```

```
                     /organism="Pichia angusta"
                     /strain="CBS 4732"
                     /db_xref="taxon:4905"
                     /clone="BB0AA021B09"
                     /clone_lib="BB0AA"
                     /note="end : T3"

misc_feature
                     <10. .>678
                     /note="similar to P07337 ( Beta-glucosidase precursor BGLS
                     )| Kluyveromyces marxianus |"
                     /evidence=not_experimental

BASE COUNT      275 a    193 c    231 g    289 t    5 others
ORIGIN
```

```
Query Match     5.3%;  Score 158.6;  DB 17;  Length 993;
Best Local Similarity   50.4%;
Matches  462;  Conservative   2;  Mismatches  446;  Indels   7;  Gaps  3;
                                              Pred. No. 3e-12;
```

```
Qy    865 CGTTCCAGAATTGCTGTGCGAGGAGACTCCCAGCCGGGGTGCGTTCATGACGGCCTACAATCA 924
          ||||| | ||||| || || ||||||||||||||||||||| || | ||||||| | ||| ||
Db     11 CCTTTCAATTGGCCCTCAAATACTCACACGCCCAAATGTGTGATGAGCTCGTACAACAACA 70

Qy    925 TCAAAGGGCGTGTCGTGCAGGGAGAGAAACCCTAAATATCCTTGATGGGATGCTTCGAAAGGAAT 984
          |||  |||| || | || ||||| |||   || | |||| || ||| |||| ||| |  |:||
Db     71 TTAACGGAGAGCACGTTTCACCAATCCAAAAAGCTGCTCACTGGTTGTGCTGCCGGAGAGGWAT 130

Qy    985 GGGGGTTGGGAGTGGCCTAATCATGACCGACTGGTACGGCGACATACAGTACCACAGAAGCCG 1044
          ||||||||||| || |||| ||||||||||| ||||  ||||| || ||| ||||| | |||||
Db    131 GGGGGTTGGGAATGGGTTTGGGTTATGTCCGACTGGTTTGGGTGGTGTCTACTGCTACGAA 190

Qy   1045 TTGTGGCAGGCCTCGACCTCGAGATGCCCGGGAGACCTCCACGCCTTCCAGACGGAGAAACACTCA 1104
          ||||| ||| || ||| || ||| |||| |||||||||| ||||| | |||| |||||||| |
Db    191 TTGATGCAGGTTTGACCTGGAGTGTCCCGGGCGTTTCCAATCATGAGAAAAATGTGTTGATGAGCTCG 250

Qy   1105 AGTTCAAACGTCTCCAA---CGGAAAGCCCTTTATCCACGTCATTGACCAGAGGGCTAGG 1161
          ||| ||||| |   ||   |||| |||| | || |||||||| || || |||||||||
Db    251 TGTTACACGCTATCAAATGCCAGAGAGATCAGTATAGACGTGATTGATGAAGAGTTAGAA 310

Qy   1162 AAGTTCTTCAGTTCGTCAAGAAGTGTGCTGCCTCCGGAGTGACCGGAGAACGGGCCCCGAGA 1221
          |||||| || ||| | ||| ||  ||| || || ||||||  |  ||||  |||||| |||
Db    311 ACGTGCTGAACTTGGTGAAGTTTTCTATGCAAAGGCAGTTCCTGAAAATGCTCCTGAAG 370

Qy   1222 CGACTGTCAACAACACACCCCGAAACGGGCAGCTCTCCTCCGGGAAGGTTGGCACGAGGGCA 1281
          | ||| ||||| | ||||  ||||||||||| || |||  |  || | ||||| | |||||
Db    371 ATTCTAAGAATAAATGTTCCTGAGACCTCTGCTTTAGCCGGAAAAGGCTAGCCTCAGAAGCTA 430

Qy   1282 TCGTGCCTGCTGAAGAAACGAGAACAACAACGTTCTGCCCTTGAGCAAGAAGAAGAAGAGCAGGCTGA 1341
          |||||| |||||||||||||||| ||||| |||||| || | ||||||||| || |||||||
Db    431 TCGTGTTGCTGAAGAATAAGGAATTCATTTTTTGCCGTTGAAAAAGAAGAGGATAAGATTGCTG 490

Qy   1342 TTGTCGGCCCCAACGCCAGCGGCCACATACCACGGCGCGGAGGCTCTGCCGCACTCAGG 1401
          ||||| || | |||| || || | ||||| |||| | ||| | ||| |||| |||||
Db    491 TGATTGGGCCCTAAAGTTCCTAGAATCTTTGGTGGGGCCTTCGATGAATC 550

Qy   1402 CCTACTACCGCAGTCACTCCCTTTGACGGGCCTCAGCAAGCAGCGTCGAGACGCGCCCATCGT 1461
          || || || |||| ||||   | | |||| ||||| ||| |||| |  |||||||||||
Db    551 CTTATTATGCTACAAACATTTTTGATGGGTATTGGCTGATAAAATTACAGAGCCCCACTCT 610

Qy   1462 ACACCGTCGGCGCGCCTACACCACCGTTCCTCCCATTCTAGGCGAGCCAGTGGCTCACGCCCG 1521
          || || |||||| ||| || |||| | |||| |||| | || |||| | | |||||||||
Db    611 ATGCTGATGGTTGCT-CAATTGACGAAAAAACCTCGACTCTGGGCCCCCACCTTGTCTACCTTC 669

Qy   1522 ACGGCGCCTCCGGGCATGCGCTGGAGGGTCTTCAACGAGCGCCCCTGGTACCCCTACCCGCC 1581
          |||||| || ||| | ||| ||||||  || ||||||| |||||  ||| | ||||| |||
Db    670 AACGGTTCCGAAGGGTTTATTGCWAGGTTTACAACCTTCCTGGTCGAATCGGTCTGACCGTG 729

Qy   1582 AGCACATTGACGAGCTCTTCTTCACCAGACGCGACAATCGCACCTGGTCGGACTACTACCACC 1641
          |||||| | | || |||| || |||| ||  |||| | | ||||| || ||  ||||||||
Db    730 TCATTGTTGATGAGTTTGCATTGAATTCTACGGAGATTGATGTGTTTGATTACAAAAACG 789

Qy   1642 CCAAGGCGGCGAGACACGTGGTACGCC---GACATGGAGGGCACGTACACCGCCGACGAGG 1698
          ||    ||||| || ||| | |||||   ||||| || || |||||| |||| ||||||
Db    790 ACAAACTAGAACAACAACTTGTTCTATATGGAATATCGAGGGCGCCAAAGAGA 849
```

```
        388 CCATGATTACAAAAAGCTAAAGCTCTCATTCAAGCTTGGTACGGAGGTAATGAGGTTGGGC 447
            |||||| |||  || ||| ||| |||| | | || || | |||| || ||||| |  |
       2241 AACTCCATTGCCGACGTGTCTTTGGCGGACTACACACCCTCGGGCAAGCTGTGCCCTCAGC 2300

Db    448 AATGCTATTGCTGATGTGTCTTGTATGGTGACGTTATTCCAAGTGGTAAATTATTCTCTCA 507
            ||||| || | ||  |||||    ||  ||   | || ||   |  |  || | | |||
       2301 TTCCCAAGCCGCCTGCAGGAACAACCCCGCGGTTTCTCAAACTTCCACCAGCGCCGGGCGC 2360

Db    508 TGGCCAATAAAACCTTGAAGATAACCCAACTTATTTGAAACTTCAAGACTGAATTTGGCAGA 567
            | ||| |  || | | | || ||  | || |  | | |||| ||| | ||  || |||
       2361 ACGGCTGTACGGCGGCGAGCAGCGTCACGTCGGGTACGAGGTTTGCCGACAAGGAC 2420

Db    628 GTTGCAATTCCCTTCGGTTACGGGTTTATTTTTATCTTATACTGAAATTTGCTTTATTCTAATCAA 687
            || || |||  |  ||   ||||  |   | ||  ||| || || ||| ||
       2421 GTCAATTTCCCCTTTGGCCACGCGTCCTCTACCACCACTTTTGCCGTTTTCCAATCTCTCC 2480

Db    688 GTACAAACGGATAGAGAGGTTATTCTCTGTTTCTGTTAAATGTAAAGAACGCAGG 740
            || || || | | || ||||| | || | | || | || |  |  |  |||
       2481 GTGCTCACAAGGACGGGCAAGCTGAGCGTGTCCCTCTCCGTGAAGAAACACGG 2533
```

RESULT 12
CNS076GU/c

```
Qy  1699 ACTGCACCTACGAGACTCGGTCGTCGTCTCTGCGGCACGGCAAAGGGGTACGTAGACGACC 1758
         |||||||||||| |||||||||
Db   850 CTGGCACCTACGAGATAAAGCGTTAGTTGTTCTTGGACCGCACAATTTTTCATTAACGGTA 909

Qy  1759 AGCTCGTCGTCGACAAC 1775
         ||||||||||
Db   910 AATTGGAGATTGATGAC 926
```

LOCUS   CNS076GU   986 bp   DNA   linear   GSS 07-JUL-2001
DEFINITION T3 end of clone BB0A003F03 of library BB0AA from strain CBS 4732
            of Pichia angusta, genomic survey sequence.
ACCESSION   AL431380
VERSION   AL431380.1 GI:12214792
KEYWORDS   GSS.
SOURCE   Pichia angusta
  ORGANISM Pichia angusta
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1 (bases 1 to 986)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   MEDLINE   20584711
    PUBMED   11152876
REFERENCE   2 (bases 1 to 986)
  AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 13. Pichia
            angusta
  JOURNAL   FEBS Lett. 487 (1), 76-81 (2000)
   MEDLINE   20584723
    PUBMED   11152888
REFERENCE   3 (bases 1 to 986)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES   Location/Qualifiers
     source   1..986
                     /organism="Pichia angusta"
                     /strain="CBS 4732"
                     /db_xref="taxon:4905"
                     /clone="BB0A003F03"
                     /clone_lib="BB0AA"
                     /note="end : T3"
     misc_feature   complement(<27..>959)
                     /note="similar to P07337 | Beta-glucosidase precursor BGL5
                     1 | Kluyveromyces marxianus 1 "
                     /evidence=not_experimental
BASE COUNT   279 a   237 c   195 g   274 t   1 others
ORIGIN
Query Match   5.3%;   Score 158.4;   DB 17;   Length 986;

Best Local Similarity 51.6%; Conservative 0; Mismatches 381; Indels 9; Gaps 2;
Matches 416; Pred. No. 3.2e-12;

```
Qy  1515 ACGCCCGACGGCGGGGCGGTCCGGGCGCATGCGCGTCGGAGGGTGCTTCAACGAGCCCCT 1574
         |||  ||  ||||  |||||
Db   827 ACCTTCAACGGGTTCCGAAGGGGTTTATTGCCAAGGTTTACACACCTTCCTGCTGAAATCGTCT 768

Qy  1575 AACCGCACCAGCACATTGACGAGCTGCTCTTCTTCACCAAGAGGACAGGACAATGCCACCTGGTGGGACTAC 1634
         ||||
Db   767 GACCGTGTCATTGTTGATGGAGTTTGATGAAATTCTACGAGATTGATGTGTTTTGATTAC 708

Qy  1635 TACCACCCCAAGGCGGCAGACA---CGTGGTACGCGCGACATGGAGGGCACGTACACCGCC 1691
         |||
Db   707 AAAAACGGACAAACTAGACAACAACTTGTTCGTGTCTATATNGGGATATCGAGGGCGCCCC 648

Qy  1692 GACGAGGACGTGCCACCTTACGAGCTGGGCGTCGTCGTGGGCGACGGCAAAGGGCGTACGTA 1751
         ||||||||||||||||||||||||
Db   647 AAAGAGACTGGCACCTACGACGATAAGGCGTTAGTTGTTCTTGGAACCGGCACACATTTTTCATT 588

Qy  1752 GACGACGACCAGCTGCGTCGTCGACAAACCGCACAAGCAGGTCCCGGCGGCGGATGCGCTTCTCGGC 1811
         |||||
Db   587 AACGGTAAAATTGGAGATTGATGACAAAAAATAATCAAGTCCTTGGAATTTGCAGCCTGGGT 528

Qy  1812 TCCGCCACCCGCGGGAGGAGCAGGGGCCGCCATCAAATCTCGTCAAGGGCAACACGTACAAAGTTC 1871
         ||||||     |||||||||    |||
Db   527 GCCTCTACCGCCCCCAAAACCAAGAGATTTCATATGAACAAAGGGAGAGTCTTTCAAAATAC 468

Qy  1872 AAGATCGAGTTCGGCCTGGCACCCACTACACCTCAAGGGCGACACCATCGGTCCC---- 1927
         ||||||||||||||
Db   467 AAGATTGAGTTTGGTTCGGGGTCCAACGTTCACTTTGGAGACCACTGATCTCAGTTGCATCT 408

Qy  1928 --CGGCCACGGCTCCCTGCCCTCGGCGGCGTGGGGCGCCGCTCAAGGTCATTGACGACGACCAGGCCGAAATC 1985
         |||
Db   407 AACGGGAGGAGGAGCCGCTCTATGTTCCGCCCATTGAAGGTCGAGTCGGATAAGGATAGGATC 348

Qy  1986 GAAAAGTCCGTGCCCTCGCCCAAGGAGCACGACGACCAGGTCAATCATCTGCGCGGGGCCTTAAC 2045
         |||||||||||
Db   347 AAAAAAGGCTGTTGAAGTTGCATCACAAACTGATAAAGTTGTTCTCTTGTGTTGGGGAACATCA 288

Qy  2046 GCCGGACTGGGGAGGACCGGAGGGGCGCGACCGCGGGGAGCGAGCGAGCTCCCCGGCGCGTGGGTGGGAC 2105
         |||
Db   287 TCAGATTGGGGAGGTGAGGGGTACGATAGGACCAGTTATGGGACTTCCTGGTAGCCAGGAT 228

Qy  2106 CAGCTCATTGCCGACGTGGCGCCGCGGGCGGAACCCAAAACACCGTCGTCGTCGTCATGCGAGACGGGC 2165
         |||
Db   227 GAGCTCATCCAAGAGAGTCCTTAAAGTCAACAACGAAATGTGGTGGTAGTCAATCTTTTCCGGT 168

Qy  2166 ACCCCCGACGAGGAGATGCCCCTGGCGTCGACGCCACGGCGCCGCGGTCATCCAGGCCTTGGTACGGC 2225
         ||||||
Db   167 ACCCCGAGTCACGATGCCATNGGGTCGATGAGGGTTCCGGCGTTATAAATTCACGGATGGGTCAAT 108

Qy  2226 GGCAACGAGACGACGGGCGCAACTCCATTGCCGACGTGCTCTCTTGGCGGACTACAACCCCTCGGGGC 2285
         ||||  |||||||
Db   107 GGAAGTGAGAGTGGAAATGCAAATTGCTGATTGTTGTTAACCCGTCCGGT 48

Qy  2286 AAGCTTGTCCCTCAGCTTCCCCAACCG 2311
         ||||||| || ||||
Db    47 AAATTGAGTTTAACATTCCCTAAACG 22
```

RESULT 13
CNS074WR

LOCUS   CNS074WR   836 bp   DNA   linear   GSS 07-JUL-2001
DEFINITION clone BA0AB031C06 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL429361
VERSION   AL429361.1 GI:12212555
KEYWORDS   GSS.
SOURCE   Kluyveromyces lactis.
  ORGANISM Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1 (bases 1 to 836)

(Left column — Kluyveromyces lactis GSS)

```
AUTHORS    Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
           Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
           de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
           Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
           Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
           Wincker,P. and Weissenbach,J.
  TITLE    Genomic exploration of the hemiascomycetous yeasts: 1. A set of
           yeast species for molecular evolution studies.
  JOURNAL  FEBS Lett. 487 (1), 3-12 (2000)
   MEDLINE 20584711
   PUBMED  11152876
REFERENCE  3  (bases 1 to 836)
  AUTHORS  Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
           Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
           Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE    Genomic exploration of the hemiascomycetous yeasts: 11.
           Kluyveromyces lactis.
  JOURNAL  FEBS Lett. 487 (1), 66-70 (2000)
   MEDLINE 20584721
   PUBMED  11152886
COMMENT    This GSS is part of a random genomic sequencing program of thirteen
           yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
           exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
           Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
           lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
           angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
           Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
           5 kb were prepared and both extremities were sequenced. See
           keywords for description of this sequence and for the sequence of
           the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..836
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB031C06"
                     /clone_lib="BA0AB"
     misc_feature    <7..>780
                     /note="similar to P07337 [ betaglucosidase ]"
                     /evidence=not_experimental
BASE COUNT    238 a    163 c    192 g    236 t      7 others
ORIGIN
```

```
Query Match    5.2%;  Score 153.4;  DB 17;  Length 836;
Best Local Similarity  52.1%;
Matches  385;  Conservative  4;  Mismatches 340;  Indels 10;  Gaps 2;
                                  Pred. No. 1.6e-11;

Qy  1658 GTGGTACGGCCGACATCGGAGGGCACGTACACCCGCGACGAGGACTCGACCTACGAGCTCCG 1717
            ||| |||||  | |||| ||| ||  || ||  |   | || || ||| | |  |||
Db   102 GTTCTACATCACACTGGAAGGTTACTTACTCCTAAAGAAGATGGCAACTACCTTTTCGG  161

Qy  1718 CCTCGGTCGTCGTTGCGGCGACGGACAGCGTACGTAGACGACCAGCTCGTCGTCGTCGACAACGC 1777
            |||||  |||   |||||     |||  ||||  |||   ||  || || ||| |||
Db   162 TTTACTCGGTTTACGGCGACCACCGGACATATTGTACTTGGGATGATGAATTAGTGGTTGATTCAGA 221

Qy  1778 CACCAAGACAGGTCCCCGGCCGATGCCTTCTTCGGCGTCCGGCCACCCGCGAGGACGGGCCG 1837
             | |   ||| || |||| ||| ||     || ||  | || || | | ||| |||
Db   222 GAMAGACCAAAACTCGTGGAAGTTCTGCTTCGTCGGTTGCTTACCGGATGGAAAAAGACCAAGAC 281

Qy  1838 CATCAAATCTCGTCAAGGGCAAACACGTACAAGTTCAAGATCGAGTTCGGCTCCGCCACCCAC 1897
            |||| ||| || |||||  |||  | ||  |  | | || ||| ||  || || || ||
Db   282 TGTCACTCTCGCAAAAAGGGTAAAGCCTATAAAGTTAGAAACGGATTACGGTTCTGGTCCAAC 341

Qy  1898 CTACACCCTCAAGGGCGACGACCATCGTCCCCGGCCACGGCCTCCCTCCGGCGTCGGGCGCTG 1957
            | || || |||||  |||||     | | ||  || |||||  ||  | || ||||||
Db   342 AAGTCAGTTGGTCAGTGA------AATTGGATCTGGCGGGGTTTTCAAGTTGGTGTTGGTGTTGC 392
```

RESULT 14

```
LOCUS       CNS0770T/c              860 bp    DNA     linear   GSS 08-JUL-2001
DEFINITION  CNS0770T
            T3 end of clone BB0A008D11 of library BB0A from strain CBS 4732
            of Pichia angusta, genomic survey sequence.
ACCESSION   AL432099
VERSION     AL432099.1  GI:12215513
KEYWORDS    GSS.
SOURCE      Pichia angusta
  ORGANISM  Pichia angusta
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   MEDLINE  20584711
    PUBMED  11152876
REFERENCE   2  (bases 1 to 860)
  AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 13. Pichia
            angusta
  JOURNAL   FEBS Lett. 487 (1), 76-81 (2000)
   MEDLINE  20584723
    PUBMED  11152888
REFERENCE   3  (bases 1 to 860)
  TITLE     Direct Submission
  JOURNAL   Genoscope.
            Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
```

```
Qy  1958 CAAGGTCATTGACGACCAGGCCGAAATCGAAAAGTCCGTCGCCCTCGCCAAGGAGCACGA 2017
            |||||   ||| || |||||||  ||| |||||||| ||| ||| | ||||||| |||
Db   393 CAAGGCTATTGATGCAGATGAAGAAATCAGGAAGTCCGGAGACCGAAAACTTGCTGCGGGTCACGA 452

Qy  2018 CCAGGTCATCATCTGCGCGGGGGCCTTAACGCCGACNTGGGAGACCGAGGACGGGCGC 2077
            ||||| ||||| ||| |||  |||| |||| |||| | |||||||| || |||| |||
Db   453 CAAAGCCATTTTATGTATTGGTTTTGAACCGACAGAAATGGGAATCTGAAGGTCATGATAGGGA 512

Qy  2078 GAGCATGAAGCTCCCCGGCGGTGCTCGGACCAGCTCAATTGCCGACGTGCGCCCCGGCCAACCC 2137
            || |||||| || || ||||| | | ||  || |||||   ||  || ||| || ||
Db   513 AGATATGATCTTTGCCAGGMAGAACAAAAACGATCTCGTTCCGTGCTTAGAGGCCAAATCC 572

Qy  2138 AAACACCGTCGTCGTCATGCAGACAGCGGGCACCCCGCAGAGATGCCCTGGCTCGAGCGCCAC 2197
            |||| || || || || || ||| ||  || || || || || |||| ||  || |||| |||
Db   573 TAACACTGTCATCATCAACACAGTCTGGTACTCCCCGTAGAATTTCCATGGGTAGAACAAGC 632

Qy  2198 GCCCGCGTCATCCACGGCGTCTGGTACGGCCGCCAACGAGACGGGCAACTCCATTGGCGACGT 2257
            || | ||  |||| ||||||  | || ||  |||  |  ||  || || || ||| |||| ||
Db   633 CAATGCCTTACTTCAAGCTTGGTACGGCAGGTAAATGAACTGGGCAATGCTATTGCCGATGT 692

Qy  2258 CGTCTTTGGGGACTACACACCCTCGGCCAAGCGTCCCTCAGCGTTCCCAAGCGCCCTGCA 2317
            || ||| || ||  || || || || |||||| | ||| ||| | || |||| ||| ||
Db   693 CTTGTACGGGACGTTGTTCCAGGTGGTAAATTAACTCTCATGGCCATTAAAAACTTGA 752

Qy  2318 GGACAACACCCGCGGT-TTCTCAACTTCCGACCGAGGCCGCGGGCGCCGCGGTGTACGGCGAGG 2376
            |||| || ||| ||  | ||||| ||||| || || || || || | || |||||| ||||
Db   753 AGATAACCCAAACTTAATTTTGAWCTTCAAGACTGAATTTGGCAGAGTTCTCTACGGTGAGG 812

Qy  2377 ACGTCTACGTCGGGGTACAG 2395
            || |   ||| || ||||
Db   813 ATATTTTCATTGGATACAK 831
```

FEATURES   Location/Qualifiers
  source   1..860
            /organism="pichia angusta"
            /strain="CBS 4732"
            /db_xref="taxon:4905"
            /clone="BB0AA008D11"
            /clone_lib="BB0AA"
  misc_feature   complement(<3..>794)
            /note="end : T3"
            /note="similar to P07337 [Beta-glucosidase precursor BGlS
            ] [Kluyveromyces marxianus]"
            /evidence=not_experimental

BASE COUNT   246 a   193 c   159 g   261 t   1 others
ORIGIN

Query Match   4.28;   Score 126.2;   DB 17;   Length 860;
Best Local Similarity   49.5%;   Pred. No. 6.5e-08;
Matches 357;   Conservative 0;   Mismatches 358;   Indels 6;   Gaps 1;

Qy   1668 GACAATGGAGGGCACGTACACCGCCCGAGGGAGACTGCACCTACGAGCTTCGGCCTCGTCGTC 1727
           |||||| |||||| ||| | | | ||| |||||| |||||||| | |||| |||||||||
Db   722 GACATAGAGGGGGCGAATTTACCTCTGAAAAAGAATGTGTTTATGATATATACATTGTCATCG 663

Qy   1728 TGCGGCACGGCGAAAGGCGTACGTAGACGACCAGCTCGTCGTCTGCGACAACGGCCACCAAGCAG 1787
           || | || ||||| ||||| || | ||| ||| ||||| |||| |||||
Db   662 TTGGGAACANTCTTTGCTTTTATCTAAAACGGCAAAATTATTCATAGATGATAAAACAAGCCAA 603

Qy   1788 GTCCCCGGGCGAATGCCTTCTTCGGCTCCGCCACCCGCCGAGGAGCCGCCGCATCAAATCTC 1847
           ||| ||| ||| | | | ||| |||
Db   602 AAGTTGGGGATTTGCACCAATAGGTGCGAGCTCTCTCGGCAAGACCAAGTCTTTGAAAATT 543

Qy   1848 GTCAAGGGCAAACACGTACAAGTTCAAGATCGAGTTCGGGCTCCGGACCACCTACACCCTC 1907
           || ||||| || | | ||| ||||| ||| ||| |||| || || |||| |||
Db   542 AAGAAAGACGAAACAGTGAAGTTTAAAATTGAATTTGGACATCGTGCTCCTACTTTTACGTTG 483

Qy   1908 AAGGGCGCCACCATCGTCCC---------CGGCCACGGCGCTCCCCTCCCCGCGGGCCGGCTGCAAG 1961
           ||| | |||| |||| || || |||| | |||||||||||
Db   482 GAGACCCTTGATCTGGTGGGATCAAACGGCGGAGGTTCGTTGACTGTTGCAATTGGAGAA 423

Qy   1962 GTCATTGACGAGCACCAGCCCGAAAATCGAAAAAGTCCGTCGCCCTTCGCCCCAAGGAGCACCACCAG 2021
           ||||||| || | |||| ||||| || | | | |||| ||| | ||||||||
Db   422 AGGATCCCTGATGAGGTCAGAATTAAAGATGCCAGTAGACCTCGCAAAAAAGGTCGATAAG 363

Qy   2022 GTCATCATCTCGCGCGGGCCTTAACCGCCGACTGGGAGAACCCGACGAGGGCGCCGCGAGC 2081
           ||||| ||| ||||||||| | || || || | |||||| |||||| ||||
Db   362 GTCATACTTTGTTGTATTGGAACCTCCAACGAATTGGAGAGCGAGGGATTTGACCGGCCTCAC 303

Qy   2082 ATGAAGCTCCCCGGCGGTGGCTCGGACCAGCTCATTGCCGACGTGGCGGCCGGAACCCAAAC 2141
           || |||| ||| || || |||| |||| ||| || | ||| ||| |||||
Db   302 ATGGANTCTTCCAAAACTTCAAAATAAGCTTGTTGACGAAAATACTGAAAGTCAACAGGAAC 243

Qy   2142 ACCGTCGTCGTCATGCAGACGGGCACCCCGCAGGAGATGCCCTGGGCTCGACGGCCACGCCC 2201
           || | ||| | ||| || |||||| | ||| | ||||| ||| ||| ||||||||
Db   242 GTGATTCTAGTCAAATCTATCTGGACACGCCGTAACTCTCCCTTGGATAAGATAGACGTCCCT 183

Qy   2202 GCCGTCANCCAGGCCTGTACGGCGCGGCAACGACGAGCGGCAACTCCATTGCCCACGTCGTC 2261
           ||| | | | ||| ||||| | ||| || ||| || | |||||| || | |||
Db   182 GCACTTTTACACGGATGGTTTAATGGGATGGAATCAGGAAATGCTATTGCAGACGTACTA 123

Qy   2262 TTTGGGCGACTACAACCCCTCGGGCAAGCTTCCCCAAGCGCCTTGCAGGAC 2321
           | | | || ||| ||| || | | | ||| || ||| |||
Db   122 TACGGAGACGCTAATCCTTCCGGAAAGCTTTTCTCTTTCTTGCCTAGACGTTGTGAAGAC 63

Qy   2322 AACCCCGCGGTTTCTCAACTTCCGCACCGAGGCCGCCACCGAGCTGTACGGCGGAGACGTC 2381
           || | |||| | || || || | | | ||||| | ||| |||| | ||| ||||
Db   62 AATCCTTCCTTTTTGAACTTTAAATCGACAATGGTAAAGTTGTTCTATGGCCAAGACGTT 3

RESULT 15
LOCUS   CNS075M9
DEFINITION   CNS075M9   869 bp   DNA   linear   GSS 07-JUL-2001
            clone BA0AB036G05 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL430279
VERSION   AL430279.1   GI:12213473
KEYWORDS   GSS.
SOURCE   Kluyveromyces lactis.
  ORGANISM   Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1   (bases 1 to 869)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvoglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE   Genomic exploration of the hemiascomycetous yeasts; 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
  PUBMED   11152876
REFERENCE   2   (bases 1 to 869)
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE   Genomic exploration of the hemiascomycetous yeasts; 11.
            Kluyveromyces lactis
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
  MEDLINE   20584721
  PUBMED   11152886
REFERENCE   3   (bases 1 to 869)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES   Location/Qualifiers
  source   1..869
            /organism="Kluyveromyces lactis"
            /strain="CLIB 210"
            /variety="lactis"
            /db_xref="taxon:28985"
            /clone="BA0AB036G05"
            /clone_lib="BA0AB"

BASE COUNT   250 a   149 c   184 g   285 t   1 others
ORIGIN

Query Match   4.18;   Score 120.6;   DB 17;   Length 869;
Best Local Similarity   51.2%;   Pred. No. 3.6e-07;
Matches 282;   Conservative 0;   Mismatches 269;   Indels 0;   Gaps 0;

Qy   62 AATCCTTCCTTTTTGAACTTTAAATCGACAATGGTAAAGTTGTTCTATGGCCAAGACGTT 3
          ||| | | ||||| |||| | ||| | | | | || | ||| | |||| |||||||
Db   62 AATCCTTCCTTTTTGAACTTTAAATCGACAATGGTAAAGTTGTTCTATGGCCAAGACGTT 3

Qy   2382 T 2382
          |
Db   2 T 2

Qy   2209 TCCAGGCCTGGTACGGCGCGGCAACGACGAGCGGGCGAACTCCATTGCCGACGTCGTCTTTGGCG 2268

```
Db    43  TCAAATGCTTGGTTACGGATGGGAGGTAATGAGTTGGGCAAATGCTATTTGCTGATGTCTTTGTATGGTG   102
          |||  |||  ||||||||| ||  ||| ||   |||  || |||   ||  ||||| |   ||   |||

Qy  2269  ACTACAAACCCCTCGGGGCAAGCTGTCCCTCAGCTTCCCCAAGCGCCTTGCCAGGACAACCCCG  2328
          ||  |   |||| ||||||||||||    ||  || ||||| ||||||| |    ||| ||||||
Db   103  ACGTTATTCCAAGTGGTAAATTATCTCTCTCATGGCCAATAAACCTTGAAGATAACCCAA   162

Qy  2329  CGTTTCTTCAACTTCCGCACCGAGGCCGGGCGCCACGCGTGTACGGCGGAGGACGTCTACGTCG  2388
          | || ||| ||||| || |||| | |||| | | |  | ||  | ||| || | | ||  | ||
Db   163  CTTATTTGAACTTCAGACTGAATTTGGCAGAGTTCTCTACGGTGAGGATATTTTCATTG   222

Qy  2389  GGTACAGGTACTACGAGTTTGCCGACAAGGACGTCAAATTTCCCCTTTTGGGCCACGCGCCTGT  2448
          | ||||| | ||| |||  ||| ||| || ||| ||||| ||||| |  || |||||| | | |
Db   223  GATACAGATTCTTATGAGAAGTTGCAAAGAAGATTGCATTCCCCTTCGGTTACGGGTTTAT   282

Qy  2449  CCTACACCACTTTTTGCCTTTTCCAATCTTCTCCGTGTCTCACAAGACGGGCAAGCTTGAGCG  2508
          | ||  | | ||||| |||| || | | | | | |  |  || | ||| || ||  |  | ||
Db   283  CTTATACTGAATTTGCTTTATCTACAAGTACAAAACGGATGAGGAGGTTATCTCTG   342

Qy  2509  TGTCCCTCTCCGTGAAGAACACCGGCGTCCGTGCCCCGGCCCACAGGTGGCCCAGCTCTACG  2568
          || |  ||| || ||||| | || || || || ||  |  | || ||| ||| | | |||
Db   343  TTTCTGTTAATGTAAAGAACGCAGGTGAGGCATATGCCGGGATCCGAAGTCGTACAAATTT   402

Qy  2569  TCAAGCCCCTCCAAGCGGCCAAAGATTAACCGCCCCGTCAAGGGGCTTCACGGGCTTCGCAA  2628
          || |||| || || || || ||||| || | | | | ||| || ||| | ||| ||| || |
Db   403  ACATTGCCTGCGACTGAAATCATCTTTTTCTAGGCAGTCAAAGAGTTGAAGGGATTCAAGA   462

Qy  2629  AGGTCGAACTGCAGCCCGGGCGAGCGGAAGGCGGTGAACAATCGAGGAGCAGGAGAAGTACG  2688
          |||| ||| || || ||| | || || || ||  | || ||| ||||||| || || || ||
Db   463  AAATTTTTGCTTCAACCTGGGCAAAACGAAAACAGTCAAAAGTTGATTTAATCTTGAAGGATT  522

Qy  2689  TCGGCTGGCGTATTTTGATGAGGAGGGGGATCAGTGGGTGGTGTGTCGAAAAGGGTGACTANTGAGG  2748
          ||  |  |||  | |||| || || |  || ||  ||| |||| | |  || ||| |||| | ||||
Db   523  CTATTTTCCTTCTTTCCATGAAGAAGTCGGGTAAGTGGGTAGTTCAGAGAAGCTGGGTTCAGTACCAAAG  582

Qy  2749  TTATCGTGAGC  2759
          ||| || |||
Db   583  TCTTGGTAGGC  593
```

Search completed: May 4, 2003, 04:09:53
Job time : 2740.34 secs